# Logique

*méthodes pour l'informatique fondamentale*

Paul Gochet
Pascal Gribomont

HERMES

# Logique

Collection dirigée par
Mario Borillo et Frédéric Nef

Paul Gochet
Pascal Gribomont

# Logique

*Volume 1*

méthodes pour l'informatique fondamentale

HERMES

© Editions HERMES, Paris, 1990, 1991, 1998
Editions HERMES
14, rue Lantiez
75017 Paris

ISBN 2-86601-249-6

ISSN 0988-0569

---

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale, ou partielle, faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause, est illicite » (article L. 122-4).
Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

# Remerciements

Les auteurs tiennent à exprimer leur profonde gratitude à tous ceux — et ils sont nombreux — qui les ont aidés dans leur tâche, les éditeurs MM M. Borillo et F. Nef, M$^{me}$ M. Wolper-Baudinet, M$^{lles}$ L. Bouquiaux et M.J. Toussaint, MM A. Borgers, H. Breny, M. Crabbé, Ph. Delsarte, Ph. De Toffol, E. Gillet, P. Godefroid, H. Grégoire, G. Hansoul, J. Hotterbeex, G. Hulin, Ph. Lejoly, B. Marchal, P. A. de Marneffe, M. Minsoul, L. Moreau, E. Nanni, M. Niézette, J. Pieters, F. Piccirelli, J. Riche, F. Schumacker, P. Yans. Les personnes citées, dont le rôle a été décisif, ne sont pas responsables des erreurs qui pourraient subsister.

# Table des matières


**1 Les notions fondamentales de la logique : d'Aristote à Boole 17**

1.1. Introduction . . . . . 17
1.2. Origine philosophique des notions de la logique . . . . . 17
1.3. La tradition aristotélicienne . . . . . 19
1.3.1. Termes généraux et termes singuliers . . . . . 20
1.3.2. Les dimensions de la signification des termes généraux 20
1.3.3. La proposition . . . . . 22
1.3.4. Le raisonnement . . . . . 25
1.3.5. Forme et contenu . . . . . 28
1.3.6. Validité et vérité . . . . . 29
1.3.7. Validité et évidence . . . . . 29
1.3.8. Le raisonnement déductif est-il stérile ? . . . . . 30
1.4. La tradition stoïcienne . . . . . 31
1.4.1. Le conditionnel vérifonctionnel . . . . . 32
1.4.2. Les cinq schémas de base des Stoïciens . . . . . 33
1.4.3. Variables et marque-place . . . . . 34
1.4.4. Forme inférentielle et forme propositionnelle . . . . . 35
1.4.5. Proposition conditionnelle et implication . . . . . 36
1.4.6. Les vérités de logique . . . . . 37
1.4.7. Logique aristotélicienne et logique stoïcienne . . . . . 38
1.5. L'apport de Leibniz . . . . . 39
1.5.1. De la logique formelle à la logique formalisée . . . . . 39
1.5.2. La logique de l'identité . . . . . 40
1.5.3. La relation d'identité . . . . . 41

1.6. La logique de Boole . . . . . 42
1.6.1. L'algèbre de Boole-Schröder . . . . . 45
1.6.2. Le carré logique révisé . . . . . 47
1.6.3. Le traitement algébrique des syllogismes . . . . . 49
1.6.4. La méthode des diagrammes de Venn . . . . . 52
1.6.5. Intension et extension des phrases . . . . . 53
1.6.6. De l'algèbre de Boole à la logique des propositions . . 54
1.6.7. Opération et relation . . . . . 56
1.6.8. Algèbre de Boole et logique . . . . . 57
1.6.9. Evaluation de l'apport de Boole . . . . . 58

**2 Les notions fondamentales de la logique : de Frege à Gödel 61**
2.1. La logique de Frege . . . . . 61
2.1.1. Utilité d'une langue artificielle . . . . . 61
2.1.2. Les propriétés d'une langue idéale . . . . . 63
2.1.3. Assertion et contenu propositionnel . . . . . 65
2.1.4. Fonction propositionnelle et proposition singulière . . . 66
2.1.5. Fonction propositionnelle et proposition générale . . . 67
2.1.6. L'introduction du quantificateur . . . . . 70
2.1.7. L'algèbre de Boole et l'idéographie de Frege . . . . . 72
2.1.8. L'univers du discours de la logique . . . . . 74
2.2. Notions de syntaxe logique et de sémantique formelle . . . . . 76
2.2.1. L'avènement de la théorie des modèles . . . . . 76
2.2.2. La distinction entre langage-objet et métalangage . . . 77
2.2.3. Tableau récapitulatif . . . . . 79
2.3. Contraintes associées à la calculabilité . . . . . 79
2.3.1. La notion d'effectivité . . . . . 79
2.3.2. La notion de théorème est-elle effective ? . . . . . 81
2.3.3. De l'effectivité à la récursivité . . . . . 81
2.3.4. Fonctions récursives primitives . . . . . 83
2.3.5. Ensembles récursivement énumérables . . . . . 84
2.4. Théorie de la calculabilité et métalogique . . . . . 85
2.4.1. La codification numérique du langage . . . . . 85
2.4.2. Récursivité et complétude . . . . . 86
2.4.3. Enumérabilité récursive et complétude . . . . . 88
2.4.4. Calculabilité, conséquence et déductibilité . . . . . 89
2.4.5. Décidabilité et complétude . . . . . 89
2.4.6. Algorithmes intraitables et algorithmes performants . . 90

**3 Le calcul des propositions : méthodes synthétiques 93**
3.1. Introduction . . . . . 93
3.1.1. Connecteurs naturels et connecteurs formels . . . . . 95
3.1.2. Les fonctions de vérité . . . . . 96

3.1.3. La bivalence . . . . . 96
3.1.4. La méthode des tableaux matriciels . . . . . 99
3.1.5. Raisonnements et propositions . . . . . 100
3.1.6. Les inconvénients de la méthode des tables de vérité . 100
3.1.7. L'algorithme de Quine . . . . . 100
3.2. Les méthodes déductives pour le calcul des propositions . . . . 101
3.3. Une langue formelle pour la logique des propositions . . . . . 103
3.3.1. La forme logique d'un raisonnement . . . . . 103
3.3.2. La langue formelle . . . . . 105
3.3.3. La notion de constituant . . . . . 106
3.4. La méthode axiomatique . . . . . 107
3.4.1. Axiomes et évidence . . . . . 107
3.4.2. Le calcul des propositions axiomatisé . . . . . 110
3.4.3. Le calcul des propositions interprété . . . . . 111
3.4.4. Logique des propositions ou des valeurs de vérité ? . . 111
3.4.5. Le système axiomatique de Hilbert et Ackermann . . . 112
3.4.6. Règles d'inférence . . . . . 114
3.4.7. Le fragment implicationnel du calcul des propositions . 116
3.4.8. Le calcul des propositions de Whitehead et Russell . . 118
3.4.9. Métathéorème de l'échange . . . . . 120
3.4.10. Application . . . . . 123
3.4.11. La preuve à partir d'hypothèses . . . . . 126
3.4.12. Exemple de preuve à partir d'hypothèses . . . . . 128
3.4.13. Hypothèses versus axiomes . . . . . 129
3.4.14. Axiomes et schémas d'axiomes . . . . . 130
3.4.15. Preuve à partir d'hypothèses auxiliaires . . . . . 131
3.4.16. Le métathéorème de la déduction . . . . . 132
3.5. La méthode de la déduction naturelle . . . . . 134
3.5.1. Fragment implicationnel en déduction naturelle . . . . 134
3.5.2. Exemple de démonstration en déduction naturelle . . . 135
3.5.3. Extension de la déduction naturelle . . . . . 137
3.5.4. Exemple de dérivation et de démonstration . . . . . 137
3.5.5. Règles facultatives en déduction naturelle . . . . . 138
3.5.6. Application des règles facultatives . . . . . 141
3.5.7. Méthode axiomatique et déduction naturelle . . . . . 142
3.5.8. Définition récursive de la notion de dérivation . . . . . 143
3.6. Evidence, sens et notation . . . . . 145
3.6.1. L'évidence des preuves en déduction naturelle . . . . . 145
3.6.2. Signification des constantes logiques . . . . . 147
3.6.3. Variante notationnelle utile de la déduction naturelle . 149

**4 Le calcul des propositions : méthodes analytiques 155**
4.1. La notion de séquent . . . . . 155

4.1.1. Antécédents conjoints, succédents disjoints . . . . . 157
4.1.2. Le calcul des séquents . . . . . 157
4.1.3. La démonstration dans le calcul des séquents . . . . . 160
4.1.4. Réversibilité des règles . . . . . 160
4.1.5. Démonstration de la réversibilité d'une règle . . . . . 162
4.1.6. Méthode pour démontrer un séquent . . . . . 163
4.1.7. Exemple de démonstration de séquent . . . . . 164
4.1.8. La part du choix dans la preuve en calcul des séquents 164
4.1.9. Le calcul des séquents avec coupure . . . . . 165
4.1.10. Le principe de la sous-formule . . . . . 166
4.1.11. Complétude du calcul propositionnel des séquents . . . 166
4.1.12. La portée philosophique du théorème de Gentzen . . . 167
4.1.13. Calcul des séquents et déduction naturelle . . . . . 168
4.1.14. Les constantes logiques dans le calcul des séquents . . 168
4.1.15. La méthode des tableaux sémantiques . . . . . 170
4.1.16. Le tableau sémantique en forme d'arbre. . . . . 173
4.1.17. Construction d'arbres pour un ensemble fini . . . . . 174
4.1.18. Exemple . . . . . 175
4.1.19. Une variante notationnelle de la méthode des tableaux 175
4.1.20. Tableaux *versus* séquents . . . . . 179
4.2. La méthode des connexions de Bibel et Wallen . . . . . 179
4.2.1. Complexité de la méthode des séquents . . . . . 179
4.2.2. Méthode orientée vers les connexions . . . . . 181
4.2.3. La codification de l'arbre syntaxique . . . . . 182
4.2.4. Définition de la polarité d'une formule . . . . . 183
4.2.5. Règles de prolongation et de ramification . . . . . 184
4.2.6. Types de formules . . . . . 184
4.2.7. Arbre de formule indexé . . . . . 185
4.2.8. La notion de chemin . . . . . 185
4.2.9. Evaluation de la méthode des connexions . . . . . 186
4.2.10. Application de la méthode des connexions . . . . . 187
4.3. La méthode de réduction à la forme normale . . . . . 187
4.4. La méthode de résolution . . . . . 193
4.4.1. Algorithmes pour l'homme et pour la machine . . . . . 193
4.4.2. La notion de clause . . . . . 193
4.4.3. La règle de résolution . . . . . 193
4.4.4. La procédure de résolution par saturation de niveau . 194
4.4.5. Coût de la méthode en espace de recherche . . . . . 195

**5 Le calcul des prédicats : méthodes synthétiques 197**
5.1. Introduction . . . . . 197
5.2. Syntaxe du calcul des prédicats . . . . . 197
5.2.1. Lexique . . . . . 199

5.2.2. Grammaire . . . . . 200
5.2.3. Portée et liaison d'un quantificateur . . . . . 201
5.3. L'axiomatique du calcul des prédicats . . . . . 203
5.3.1. Les nouveaux axiomes . . . . . 203
5.3.2. La définition du quantificateur existentiel . . . . . 204
5.3.3. Les règles d'inférence du calcul des prédicats . . . . . 204
5.3.4. La règle de généralisation universelle . . . . . 204
5.3.5. Définitions . . . . . 207
5.3.6. Théorème de la déduction et calcul des prédicats . . . 207
5.3.7. Métathéorème de l'échange . . . . . 209
5.3.8. Règles dérivées . . . . . 211
5.3.9. Lois principales du calcul des prédicats . . . . . 212
5.3.10. Applications . . . . . 215
5.3.11. Raisonnement logique et démonstration géométrique . 218
5.4. La méthode de déduction naturelle . . . . . 221
5.4.1. Introduction . . . . . 221
5.4.2. Introduction et élimination de quantificateurs . . . . . 221
5.4.3. La généralisation universelle dans le raisonnement informel . . . . . 222
5.4.4. L'usage de règles non valides . . . . . 224
5.4.5. Les restrictions imposées aux règles G.U. et I.E. . . . . 225
5.4.6. Justification de I.E. et de G.U. . . . . 227
5.4.7. Un regard philosophique sur la correction . . . . . 231
5.4.8. Applications de la méthode de déduction naturelle . . 232

**6 Le calcul des prédicats : méthodes analytiques 237**
6.1. La méthode des tableaux sémantiques . . . . . 238
6.1.1. Constituants syntaxiques et constituants logiques . . . 238
6.1.2. Les règles de décomposition d'une universelle . . . . . 239
6.1.3. La notion d'exemplification . . . . . 239
6.1.4. Décomposition d'une particulière . . . . . 240
6.1.5. Tableau récapitulatif des règles . . . . . 241
6.1.6. Propositions passives et propositions actives . . . . . 241
6.1.7. Nécessité d'un plan de travail . . . . . 243
6.1.8. L'univers auto-généré . . . . . 245
6.1.9. L'algorithme de construction de tableaux . . . . . 245
6.1.10. Exemple d'arbre fini fermé . . . . . 246
6.1.11. Exemple d'arbre fini ouvert . . . . . 246
6.1.12. Exemple d'arbre infini ouvert . . . . . 248
6.2. La méthode des séquents . . . . . 248
6.2.1. Les règles gouvernant les quantificateurs . . . . . 248
6.2.2. Plan de travail d'une preuve par séquents . . . . . 250
6.2.3. Intérêt de la méthode de la preuve par séquents . . . . 252

6.2.4. Complexité des preuves par séquents . . . . . 254
6.2.5. Séquents infinis . . . . . 255
6.2.6. Séquents contenant des foncteurs . . . . . 256
6.3. Matrices de Bibel et de Wallen . . . . . 256
6.3.1. Non-déterminisme de la preuve par séquents . . . . . 256
6.3.2. Calcul des prédicats et connexions . . . . . 257
6.3.3. Le chemin en calcul des prédicats . . . . . 258
6.3.4. La notion de muliplicité . . . . . 258
6.3.5. Comment définir l' inédit ? . . . . . 259
6.3.6. Alignement des universelles sur les existentielles . . . . 259
6.3.7. L'ordre de réduction des formules . . . . . 260
6.3.8. Les étapes d'une preuve par la méthode des connexions 261
6.3.9. Exemple d'une preuve par la méthode des connexions . 262
6.3.10. Les mérites de la méthode des connexions . . . . . 265
6.3.11. Application . . . . . 266
6.4. La méthode de résolution avec unification . . . . . 272
6.4.1. La démonstration automatique . . . . . 272
6.4.2. La réduction à la forme clausale . . . . . 273
6.4.3. Définitions et lemmes justifiant la réduction . . . . . 273
6.4.4. Exemple de réduction . . . . . 277
6.4.5. Construction du lexique d'instantiation . . . . . 277
6.4.6. Le théorème de Herbrand . . . . . 278
6.4.7. La méthode de résolution . . . . . 279
6.4.8. L'unification . . . . . 280
6.4.9. Contraintes sur l'unification. . . . . 282
6.4.10. Algorithmes d'unification et de résolution . . . . . 283
6.4.11. Application . . . . . 285
6.4.12. Le fondement de la méthode de Robinson . . . . . 286
6.4.13. Les limites de la méthode de résolution . . . . . 287
6.4.14. Unification et méthode axiomatique . . . . . 288
6.5. Pour clore . . . . . 289

**7 La sémantique du calcul des prédicats 293**
7.1. Deux sortes de formalismes . . . . . 293
7.2. Les ingrédients d'une sémantique . . . . . 294
7.3. Règles sémantiques pour les termes . . . . . 296
7.3.1. La notion de satisfaction . . . . . 298
7.3.2. Application . . . . . 301
7.4. Correction du calcul des prédicats . . . . . 304
7.4.1. Validité des axiomes propres au calcul des prédicats . . 304
7.4.2. Règle dérivée . . . . . 305
7.5. La vérité chez Aristote et chez Tarski . . . . . 305
7.6. Vérité et signification . . . . . 308

7.6.1. Axiomes, modèles et contenu d'information . . . . . 309
7.6.2. Catégoricité et vérité . . . . . 311
7.7. La conséquence logique . . . . . 314
7.7.1. La conséquence logique comme relation . . . . . 314
7.7.2. Trois sens du mot "complétude" . . . . . 318
7.7.3. Calcul des prédicats et théorie du premier ordre . . . . 320
7.7.4. Validité et conséquence logique . . . . . 322
7.7.5. Axiomatiser la relation de conséquence logique . . . . 324
7.7.6. Retour à la dérivabilité . . . . . 326
7.8. La nécessité du raisonnement logique . . . . . 328
7.8.1. La conception formelle de la validité . . . . . 331
7.9. Vérité et preuve . . . . . 333
7.9.1. Conditions d'adéquation . . . . . 333
7.9.2. Le paradoxe du menteur . . . . . 334
7.9.3. Solution du paradoxe du menteur . . . . . 334
7.9.4. Métalangage sémantique et métalangage syntaxique . . 335
7.9.5. Ecart entre les notions de vérité et de démontrabilité . 335

**8 Théorie de la preuve** **337**
8.1. La crise des fondements des mathématiques . . . . . 337
8.2. Les théorèmes de Löwenheim et de Skolem . . . . . 340
8.3. Le théorème de Herbrand . . . . . 341
8.3.1. Le passage de l'infini au fini chez Herbrand . . . . . 341
8.3.2. Forme clausale . . . . . 342
8.3.3. Modèle de Herbrand . . . . . 342
8.3.4. Correspondance entre interprétations . . . . . 344
8.3.5. Inconsistance sémantique dans le domaine de Herbrand 346
8.3.6. La notion de clause fondamentale . . . . . 346
8.3.7. Le théorème proprement dit . . . . . 347
8.3.8. Les arbres sémantiques . . . . . 348
8.3.9. Construction de l'arbre sémantique complet . . . . . 351
8.3.10. Les quatre étapes de la démonstration de Herbrand . . 355
8.3.11. Lemme de König . . . . . 355
8.3.12. Démonstration détaillée du théorème de Herbrand . . 355
8.3.13. Le constructivisme de Herbrand . . . . . 356
8.3.14. De Herbrand à Henkin . . . . . 357
8.4. Complétude pour les séquents du premier ordre . . . . . 359
8.4.1. Préliminaire . . . . . 359
8.4.2. Lemme . . . . . 360
8.4.3. Exemple . . . . . 362
8.4.4. Construction quasi-mécanique . . . . . 363
8.5. Problème de l'arrêt et problème de correspondance . . . . . 363
8.5.1. Problèmes et instances de problèmes . . . . . 365

8.5.2. Machines de Post . . . . . 365
8.5.3. D'un problème à l'autre . . . . . 367
8.5.4. La réduction . . . . . 369
8.5.5. Réduction de l'arrêt à la correspondance . . . . . 369
8.5.6. Conversion des instructions . . . . . 370
8.5.7. Exemple de conversion . . . . . 370
8.6. Semi-décidabilité de la logique du premier ordre . . . . . 373
8.6.1. Théorème de Floyd . . . . . 373
8.6.2. Théorème de Church . . . . . 375
8.7. Du calcul des prédicats aux théories du premier ordre . . . . . 375
8.8. La ligne de démarcation entre logique et mathématique . . . . 376

**9 Théories du premier ordre** **379**
9.1. Introduction . . . . . 379
9.2. Quelques définitions . . . . . 382
9.2.1. Généralités . . . . . 382
9.2.2. Résultats élémentaires de la théorie des groupes . . . . 385
9.2.3. Applications . . . . . 388
9.3. Théorie de l'ordre . . . . . 391
9.3.1. Ordre partiel . . . . . 391
9.3.2. Ordre total . . . . . 392
9.3.3. Bon ordre . . . . . 402
9.3.4. Le principe d'induction . . . . . 403
9.3.5. Applications . . . . . 407
9.4. Théorie des modèles . . . . . 410
9.4.1. Deux résultats essentiels . . . . . 410
9.4.2. Relations entre théories . . . . . 415
9.4.3. Relations entre structures . . . . . 417
9.4.4. Classification des théories . . . . . 421
9.4.5. La logique propositionnelle . . . . . 422
9.4.6. Applications . . . . . 423
9.5. Décidabilité . . . . . 429
9.6. Structures algébriques usuelles . . . . . 431

**10 Bibliographie** **435**

**11 Index** **447**

# Avant-propos

Les trois citations ci-dessous résument mieux qu'un long discours la part croissante de la logique en informatique :

> Il est raisonnable d'espérer que la relation entre le calcul (*computation*) et la logique mathématique sera aussi fructueuse ... que celle qui s'est installée entre l'analyse mathématique et la physique au cours du siècle dernier.
>
> John Mc Carthy

> Au cours de la plus grande partie de ce siècle, la logique mathématique a été principalement utilisée pour l'introspection. Comme outil pour la création de preuve dans la pratique quotidienne, on ne lui a guère donné sa chance. Pour qu'elle puisse réaliser ses potentialités, il semble essentiel de concevoir le but de la logique comme étant non de mimer le raisonnement humain, mais de fournir un substitut à ce dernier qui soit de la nature d'un calcul.
>
> Edsger Dijkstra

> Les connexions entre la logique et l'informatique s'accroissent rapidement et s'approfondissent. A côté de la démonstration automatique, de la programmation logique, de la spécification et vérification de programme, d'autres secteurs révèlent une fascinante interaction mutuelle avec la logique, qu'il s'agisse de la théorie des types, de la théorie du

parallélisme (*concurrency*), de l'intelligence artificielle, de la théorie de la complexité, des bases de données, de la sémantique opérationnelle et des techniques de compilation.

José Meseguer

L'ouvrage présente l'ensemble des méthodes de preuve de la logique des prédicats — des plus classiques aux plus récentes — en montrant leurs avantages et leurs inconvénients respectifs. Des applications nombreuses et détaillées, notamment aux théories du premier ordre, et des exercices de complexité variée illustrent ces méthodes.

Les concepts fondamentaux de la logique sont replacés succinctement dans le contexte des problèmes philosophiques qui leur ont donné naissance et qui, rétrospectivement, nous aident à les comprendre. Aucune concession n'est faite à l'érudition gratuite. Les recours à l'histoire y ont un rôle purement didactique.

La longueur du tome 1 — qui peut surprendre s'agissant d'un tome introductif — est due à la volonté des auteurs de désamorcer systématiquement les nombreux pièges que la logique recèle pour le novice et d'illustrer par des exemples traités exhaustivement les méthodes que celui-ci devra apprendre à manier avant de s'essayer à des applications non triviales.

# Chapitre 1
# Les notions fondamentales de la logique : d'Aristote à Boole

## 1.1. Introduction

Cet ouvrage est destiné à la fois aux informaticiens et aux philosophes. On pourrait penser qu'en nous adressant à un auditoire apparemment hybride, nous risquons soit d'être superficiels, soit de traiter de sujets qui, tour à tour, seront non pertinents pour l'une ou l'autre de ces deux catégories de lecteurs.

Cette inquiétude sur la viabilité de notre programme n'est pas justifiée. Les préoccupations théoriques pour le *raisonnement mécanisable* sont bien antérieures aux réalisations concrètes dans ce domaine. Les fondateurs de la logique moderne (Leibniz, Boole, Peirce, Frege, Herbrand) ont élaboré tout un corps de doctrine et de méthodes dans lequel les créateurs de cette nouvelle discipline qu'on allait appeler "informatique" ont puisé largement.

Par ailleurs, il se trouve que les conditions très astreignantes que les protagonistes de la *logique formalisée* se sont imposées dans leur quête de rigueur coïncident largement avec les exigences que l'on doit satisfaire dans l'élaboration d'un programme exécutable sur ordinateur.

## 1.2. Origine philosophique des notions de la logique

La logique dite "de Port-Royal" qui parut anonymement en 1662 et dont les auteurs étaient Antoine Arnauld et Pierre Nicole, s'intitulait *La logique ou l'art de penser*. On ne dirait plus aujourd'hui de la logique qu'elle est un art. On lui reconnaît universellement le statut de science, ce qui ne doit pas nous conduire à penser que les contributions à la logique des philosophes de l'Antiquité et du Moyen-Age appartiennent à un stade pré-scientifique de cette discipline.

Un grand nombre de concepts, de principes et de problèmes formulés par Aristote, Sextus, Chrysippe, Occam ou Buridan, ont été incorporés à la science actuelle. S'il est incontestable que le renouveau de la logique au XIX$^{e}$ siècle a été l'oeuvre de mathématiciens plutôt que de philosophes, il ne faut pas en conclure que la logique mathématique d'aujourd'hui n'a pas gardé une relation privilégiée avec la philosophie.

Dans un article récent intitulé "Truth of a proposition, evidence of a judgement, validity of a proof", le logicien-mathématicien P. Martin-Löf, dont les travaux sont bien connus des informaticiens, constate qu'on ne peut exposer la logique sans utiliser au moins cinq notions primitives :

- la notion de *proposition*,
- la notion de *vérité* d'une proposition,
- la notion d'*assertion* ou de *jugement*,
- la notion d'*évidence* ou de *preuve* d'un jugement,
- la notion de *correction* ou de *validité* d'une preuve.

Comment faire, se demande Martin-Löf, pour accéder à la compréhension de ces notions primitives ?

> [...] Vous devez, répond-il, vous engager dans quelque chose pour quoi vous n'êtes pas du tout préparé en tant que logicien-mathématicien [...] vous devez vous engager dans une entreprise qui est essentiellement philosophique [...] par sa nature (Martin-Löf, 1987, p. 408).

Le problème n'est-il pas simplement déplacé ? Comment la philosophie peut-elle affronter le problème du commencement radical ? N'allons-nous pas nous engager dans des controverses stériles dont le technicien et le praticien n'ont que faire ? La meilleure réponse à ces questions, on la trouvera dans l'essai *De la réforme de l'entendement* de Spinoza qui remonte à 1661 :

> Car, pour forger, il faut un marteau, et pour avoir un marteau, il faut le fabriquer. Ce pourquoi on a besoin d'un autre marteau et d'autres outils, et pour les posséder, il faut encore d'autres instruments, et ainsi infiniment. Et par ce raisonnement il serait vain d'essayer de prouver que les hommes n'ont aucun moyen de forger. Mais, au début, les hommes, avec des instruments naturels, réussirent certains objets très faciles, bien qu'avec difficulté et imparfaitement, et ceux-ci fabriqués, ils confectionnèrent d'autres objets plus difficiles avec moins de travail et plus de perfection, et ainsi, graduellement des travaux les plus simples aux outils, des outils à d'autres travaux et d'autres outils, ils arrivèrent à exécuter de nombreux et très difficiles ouvrages et sans beaucoup de travail (Spinoza, *Oeuvres complètes*, Paris, La Pléiade, 1954, p. 167).

Les auteurs de manuels de logique introduisent généralement les notions-clefs à l'aide de définitions, sans se soucier des théories dans lesquelles ces notions sont apparues. Cette vue anhistorique présente de graves inconvénients. En

isolant les notions des théories dans lesquelles elles ont été introduites, on se coupe du contexte qui leur donnait sens et des problèmes qui ont motivé la création des théories dont ces notions sont des fragments. On comprend mieux la notion de classe vide, pour nous borner à un exemple, quand on voit non seulement les problèmes que son introduction permet de résoudre, mais aussi les obstacles qu'il a fallu surmonter pour l'introduire.

Il y a un développement naturel des concepts comme il y a une évolution naturelle des espèces en ce sens que certaines idées peuvent venir trop tôt. Ainsi l'idée de créer une langue symbolique idéale apparaît dès le Moyen-Age avec Raymond Lulle, mais elle ne peut déboucher sur rien à l'époque. La raison, c'est que "Malheureusement le niveau de la science, et en particulier, celui des mathématiques, était trop bas à l'époque — il n'y avait pas encore d'algèbre — et l'entreprise échoua" (Murawski, 1988–1989, p. 66). Tantôt une idée excellente n'arrive pas à s'épanouir parce que le "terrain" approprié fait défaut, tantôt parce qu'elle se heurte à des obstacles. Ces derniers ne sont pas nécessairement des préjugés d'une époque révolue dont nous pouvons orgueilleusement nous croire préservés. Il y a des obstacles naturels — des difficultés intrinsèques — sur le chemin de la connaissance.

Dans les deux chapitres introductifs, nous supposerons, à titre d'hypothèse de travail, que la compréhension des notions de la logique exige que l'on repasse par les étapes du cheminement intellectuel au cours duquel elles ont été introduites. Ce qui revient à supposer que la loi de Haeckel selon laquelle "l'ontogénèse est une brève et rapide récapitulation de la phylogénèse" (1900) vaut aussi dans le domaine des idées.

## 1.3. La tradition aristotélicienne

La tradition aristotélicienne divise la logique en trois parties : l'étude de la conception, du jugement et du raisonnement (Eaton, 1931, p. 71). Concevoir, juger et raisonner sont trois activités de l'esprit. La manière la plus simple d'en aborder l'étude, c'est de partir de la manifestation extérieure et tangible de ces opérations mentales, à savoir de leur manifestation linguistique (comme Aristote l'a fait). Les concepts s'expriment dans le langage par des *termes généraux*. Les jugements s'expriment dans le langage par des *propositions*[1]. Les *raisonnements* s'expriment par des *suites de propositions*.

[1] Nous entendons ici par "proposition" les *phrases déclaratives* utilisées pour faire des énoncés plutôt que pour poser indirectement une question ou pour donner indirectement un ordre. Le *commandement* "Tu ne tueras pas" n'est pas une proposition. Plus loin, le mot "proposition" sera utilisé pour désigner le *sens* des phrases déclaratives.

### 1.3.1. *Termes généraux et termes singuliers*

Dans son traité *De l'Interprétation*, Aristote définit ainsi le contraste entre termes généraux et termes singuliers :

> J'appelle *universel* [général] ce dont la nature est d'être affirmé de plusieurs sujets, et *singulier* ce qui ne le peut : par exemple, *homme* est un terme universel, et *Callias* un terme individuel (Aristote, 17b, p. 87).

Le contraste entre termes singuliers et termes généraux a reçu sa formulation la plus claire chez Quine dans *le Mot et la Chose* : Quine souligne que la différence ne tient pas au nombre d'objets auxquels les termes singuliers et généraux s'appliquent respectivement, mais au nombre d'objets auxquels ils *prétendent* s'appliquer :

> [...] le nom propre "Pégase" est singulier en ceci qu'il *prétend* renvoyer à tout juste un objet, et la périphrase descriptive "satellite naturel de la terre" est générale en ceci qu'elle ne *prétend* pas, en vertu de son sens, ne s'appliquer qu'à un seul objet (Quine, 1960, 1978, pp. 147–148).

Ajoutons que la périphrase descriptive "*le* satellite naturel de la terre" est un terme singulier.

### 1.3.2. *Les dimensions de la signification des termes généraux*

La relation exprimée par le verbe "représenter" dans le contexte de la phrase "les traits de la carte en gras représentent les autoroutes" est une relation binaire. On pourrait penser que la relation exprimée par le verbe "signifier" est aussi une relation binaire unissant les signifiants (ici des termes généraux) aux signifiés.

Cette conception de la signification des termes généraux est trop simple. Certes, on pourrait dire que les termes généraux comme "homme" signifient un concept, mais on doit aussitôt distinguer *deux* aspects de celui-ci : l'intension et l'extension. Dans le *Dictionary of Philosophy* de D. Runes (1960), A. Church définit ainsi les notions classiques d'*intension* et d'*extension* d'un concept :

> L'*intension* d'un concept consiste dans les qualités ou propriétés qui constituent le concept. L'*extension* d'un concept consiste dans les choses qui tombent sous le concept (Runes, 1960, pp. 147–148).

La notion d'extension ne fait apparemment pas de problème. On dit généralement que l'extension d'un terme général, c'est la *classe* des êtres auxquels il s'applique.

Dans *Understanding Natural Language* (1972), T. Winograd décrit un dialogue entre un agent humain et un système programmé contrôlant une caméra de télévision et un bras articulé. Des messages sont communiqués par imprimantes. Ce système répond à des questions et exécute des actions dans

une scène qui comporte des cubes et des pyramides de différentes tailles et de différentes couleurs posés sur une table où se trouve également une boîte. Imaginons la scène suivante représentée à la figure 1.1.

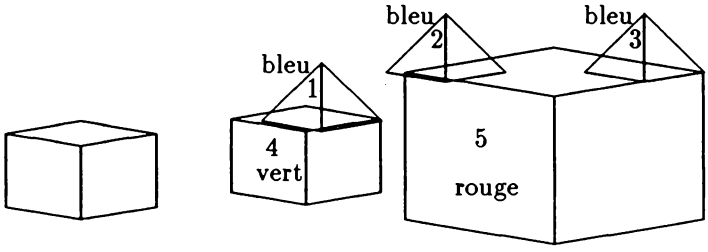



**Figure 1.1** *Blocs de Winograd*

Les termes généraux "pyramidal" et "bleu" ont *même extension*. En effet la classe des objets bleus et la classe des objets pyramidaux coïncident : ces deux classes ont les mêmes membres. En revanche ils diffèrent par l'*intension* : être pyramidal et être bleu sont des propriétés différentes. De même les termes généraux "placé dans le coin de" et "surplomber" ont *même extension*, à savoir les couples ordonnés d'individus suivants : ⟨pyramide 1, cube 4⟩, ⟨pyramide 2, cube 5⟩, ⟨pyramide 3, cube 5⟩, mais ils ont une *intension différente* : "placé dans le coin de" et "surplomber" sont des relations différentes : on peut demander au robot de déplacer les pyramides de manière telle qu'elles continuent d'être en surplomb, mais qu'elles cessent d'être dans le coin. Lorsque l'ordre aura été exécuté, l'extension de "placé dans le coin de" sera un *ensemble vide* de couples ordonnés, tandis que l'extension de "surplomber" sera restée intacte : elle sera toujours formée des trois couples ordonnés initiaux.

On parle aussi de définition d'un *ensemble* en *extension* ou en *intension* : un ensemble est défini en extension si on énumère ses membres. Il est défini en intension s'il est défini à l'aide d'une propriété que possèdent tous ses membres et eux seuls. Par exemple, la périphrase "les nombres entiers positifs ou nuls multiples de 3 et plus petits que 20", que l'on note

$$\{x : x \in \mathbf{N} \wedge 3 \mid x < 20\},$$

définit en intension l'ensemble qui est défini en extension par {0, 3, 6, 9, 12, 15, 18}. Les accolades remplacent ici la locution "l'ensemble de ...".

### 1.3.3. *La proposition*

Platon, le premier, a isolé le concept de *proposition*. Ni une liste de verbes, ni une liste de noms ne constitue un *discours*, c'est-à-dire une *proposition*. En revanche, la combinaison d'un verbe avec un nom, comme "l'homme apprend", forme un discours auquel on peut attribuer la qualité de vrai ou de faux.

> L'ETRANGER : Par exemple, *marche*, *court*, *dort*, et tous autres verbes signifiant des actions, même à les dire tous à la file, n'en forment pas plus, pour cela, un discours [...] Que donc, une autre fois, l'on dise : *lion*, *cerf*, *cheval*, et tous autres noms qui dénomment les sujets faisant les actions, c'est là encore une suite d'où n'est jamais résulté aucun discours [...] Quand on dit : *l'homme apprend*, il y a là, pour toi, discours, le plus simple et le premier ?
> THÉÉTÈTE : Pour moi, oui (Platon, *Le Sophiste*, 262, b, c).
> L'étranger fait ensuite découvrir que les discours peuvent être dits vrais ou faux.
> THÉÉTÈTE : L'un [le discours "Théétète vole"]...est faux; l'autre [le discours "Théétète est assis"], vrai (*Ibid.*, 263, b).

Les termes sont les constituants de la proposition. Ils ont une signification, mais ils ne suffisent pas à affirmer ou nier quelque chose :

> [...] le mot *homme* signifie bien quelque chose, mais non pas cependant qu'il est ou n'est pas : il n'y aura affirmation ou négation que si on ajoute quelque chose (Aristote, *De l'Interprétation*, 16b30).

Aristote a défini ainsi le concept de proposition :

> [...] tout discours n'est pas une proposition, mais seulement le discours dans lequel réside le vrai ou le faux, ce qui n'arrive pas dans tous les cas : ainsi la prière est un discours, mais elle n'est ni vraie ni fausse (*Ibid.*, 17a5).

Pour Aristote, il n'y a de science que du général. La zoologie, par exemple, traite de l'espèce cheval mais pas de Bucéphale, le cheval d'Alexandre. Cette conception philosophique de la science retentit sur sa logique. Celle-ci accorde une place privilégiée aux propositions qui ne contiennent que des termes généraux par opposition à celles qui comportent un terme singulier comme "Socrate est blanc".

Les propositions qui interviennent dans la théorie du raisonnement chez Aristote sont des propositions générales dans lesquelles figurent deux termes généraux reliés par le verbe copule "est" ou "n'est pas". Il est traditionnel de les regrouper dans le carré logique représenté à la figure 1.2 :

Aristote décèle entre ces propositions des rapports logiques qui sous-tendent des raisonnements valides très élémentaires qu'on appelle *inférences immédiates*. Par exemple, entre les propositions $A$ et $E$ existe un rapport de *contrariété* mutuelle. Ces deux propositions ne peuvent être vraies ensemble, quoiqu'elles

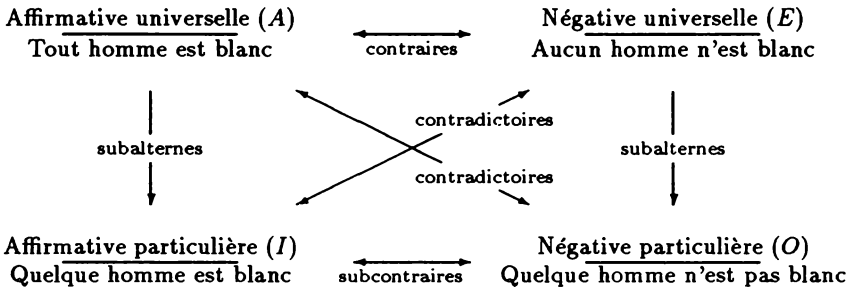



**Figure 1.2** *Carré logique*

puissent être fausses ensemble. Dès lors si une proposition $A$ est vraie, on peut conclure que la proposition $E$ correspondante est fausse. Les *contradictoires* ne peuvent être ni vraies ni fausses ensemble, les *subcontraires* ne peuvent être fausses ensemble.

L'existence de quatre sortes de propositions, ni plus ni moins, n'est pas un accident : il y a exactement quatre relations possibles entre les extensions de deux termes généraux. Le mathématicien Euler a représenté ces quatre relations par des diagrammes géométriques dans ses *Lettres à une Princesse d'Allemagne* (1761).

L'inclusion totale de l'extension du terme sujet $S$ dans l'extension du terme prédicat $P$, est représentée à la figure 1.3. C'est le cas de la proposition $A$.

**Figure 1.3** *Inclusion totale*

L'exclusion totale de l'extension du terme sujet et de l'extension du terme prédicat est représentée à la figure 1.4 (proposition $E$).

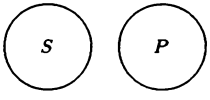



**Figure 1.4** *Exclusion totale*

L'inclusion partielle de l'extension du terme sujet dans l'extension de terme prédicat est représentée à la figure 1.5 (proposition $I$).

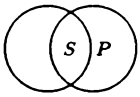



**Figure 1.5** *Inclusion partielle*

L'exclusion partielle de l'extension du terme sujet dans l'extension de terme prédicat est représentée à la figure 1.6 (proposition $O$).

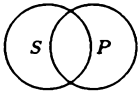



**Figure 1.6** *Exclusion partielle*

J.D. Gergonne, dans "Essai de dialectique rationnelle" (*Annales de mathématiques*) en 1816–1817, reconnaît une cinquième relation possible entre les extensions des termes : l'inclusion totale et réciproque du terme sujet dans le terme prédicat et du terme prédicat dans le terme sujet. Il s'agit d'un cas limite.

**Figure 1.7** *Inclusion totale et réciproque*

Les diagrammes d'Euler expriment les relations entre les extensions des termes généraux figurant dans les quatre propositions $A$, $E$, $I$, $O$, pour autant que ces termes généraux n'aient pas une extension vide comme c'est le cas pour le terme général "licorne" ou le terme général "centaure".

Nous verrons plus loin que G. Boole se libérera des présuppositions d'existence faites implicitement par Aristote dans son interprétation des quatre propositions fondamentales et qu'il fera place à des termes généraux dont l'extension est vide. Pour l'interprétation booléenne des propositions $A$, $E$, $I$, $O$, les diagrammes d'Euler ne conviennent plus. Il est cependant possible de représenter par des diagrammes les relations entre les extensions de termes généraux figurant dans les propositions $A$, $E$, $I$, $O$, mais il faut recourir à d'autres diagrammes. C'est à J. Venn que revient le mérite d'avoir inventé une représentation diagrammatique qui convienne aux propositions $A$, $E$, $I$, $O$ (cf. section 1.6.2) interprétées à la manière de Boole dans *Symbolic Logic* (1881). Les diagrammes de Venn nous fournissent un algorithme commode pour distinguer les syllogismes valides des syllogismes non valides dans l'interprétation

booléenne.

### 1.3.4. *Le raisonnement*

On a raisonné bien avant de faire la théorie du raisonnement. Aristote inaugure la science du raisonnement et formule la théorie du syllogisme qu'il définit ainsi :

> le *syllogisme* est un discours dans lequel, certaines choses étant posées, quelque chose d'autre que ces données en résulte nécessairement par le seul fait de ces données (Aristote, *Premiers Analytiques*, 24, b, 15).

Le syllogisme catégorique qu'étudie Aristote est un raisonnement formé de trois propositions prises dans l'ensemble des quatre propositions $A$, $E$, $I$ et $O$. Il contient trois termes généraux qui reviennent chacun deux fois comme c'est le cas dans le syllogisme 1.1.

(1.1) Tous les mammifères sont des vertébrés
Tous les hommes sont des mammifères
donc tous les hommes sont des vertébrés.

Le terme qui est sujet de la conclusion (ici "hommes") est appelé *mineur*. Le terme qui est prédicat de la conclusion (ici "vertébrés") est appelé *majeur*. Le terme qui apparaît deux fois dans les propositions dont on tire la conclusion — propositions qu'on appelle *prémisses* — porte le nom de *moyen terme*. Dans l'exemple, le moyen terme est "mammifères". La prémisse qui contient le terme majeur est *la majeure* et celle qui contient le mineur, la *mineure*. Si l'on convient d'écrire les syllogismes selon l'ordre décrit à la figure 1.8, et si l'on se

Prémisse majeure
Prémisse mineure
Conclusion

**Figure 1.8** *Ordre canonique des propositions dans un syllogisme*

souvient que majeur ($P$), mineur ($S$) et moyen ($M$) peuvent être soit sujet (à gauche de la copule), soit prédicat (à droite de la copule), on constate qu'il y a quatre figures possibles, c'est-à-dire quatre manières de placer le moyen terme dans les prémisses.

Puisque chaque prémisse ou conclusion est une proposition et qu'il y a quatre sortes de propositions ($A$, $E$, $I$, $O$), il y a $4^3$, c'est-à-dire 64 syllogismes par figure, d'où 256 syllogismes en tout. Dans le cas des syllogismes comme dans le cas des inférences immédiates, ce qui sous-tend le raisonnement, ce sont les rapports logiques existant entre les termes, c'est-à-dire les rapports d'inclusion totale, d'exclusion totale, d'inclusion partielle ou d'exclusion partielle. Tous les syllogismes ne sont pas valides. Aristote et, à sa suite, les logiciens du

| | 1ere Figure | 2e Figure | 3e Figure | 4e Figure |
|---|---|---|---|---|
| Prémisse majeure | $MP$ | $PM$ | $MP$ | $PM$ |
| Prémisse mineure | $SM$ | $SM$ | $MS$ | $MS$ |
| Conclusion | $SP$ | $SP$ | $SP$ | $SP$ |

**Figure 1.9** *Tableau des figures du syllogisme*

Moyen-Age, ont cherché non seulement à séparer les syllogismes *valides*, dans lesquels la conclusion suit nécessairement des prémisses, des syllogismes non valides dans lesquels la conclusion ne suit pas nécessairement des prémisses, mais aussi à trouver le principe qui sous-tend la répartition des syllogismes en valides et non valides.

On décèle aisément le statut (valide ou non valide) des syllogismes de la première figure en inspectant les rapports entre les extensions de leurs trois termes au moyen de diagrammes. On peut ensuite formuler des règles de conversion qui permettent de transformer les syllogismes des autres figures dans ceux de la première. Comme notre objectif n'est pas d'exposer de manière exhaustive la théorie du syllogisme, mais seulement d'introduire les notions-clefs de cette théorie et de poser les problèmes qu'elle a soulevés et qui ont conservé toute leur actualité, nous nous bornerons à examiner *deux* syllogismes à titre d'exemples.

Le syllogisme 1.2 est formé des propositions $E$, $A$, $E$ :

(1.2) Aucun poisson n'est un mammifère
Tous les brochets sont des poissons
donc aucun brochet n'est un mammifère.

On peut dégager sa forme à l'aide des *marque-place* $S$, $P$, $M$, où $S$ indique la place du mineur, $P$ du majeur et $M$ du moyen terme :

(1.3) Aucun $M$ n'est $P$
Tous les $S$ sont $M$
donc aucun $S$ n'est $P$.

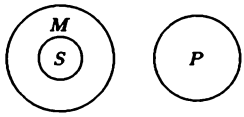



**Figure 1.10** *Syllogisme 1.4*

Cela peut s'écrire de manière plus concise compte tenu des différentes copules figurant dans les propositions $A$, $E$, $I$, $O$ :

(1.4) *MeP*
*SaM*
*SeP.*

En inspectant le diagramme, on voit tout de suite que si l'extension du terme $M$ et celle du terme $P$ s'excluent totalement et que si l'extension du terme $S$ est totalement incluse dans celle de $M$, celle de $S$ est totalement exclue de celle de $P$.

Le syllogisme 1.5 est formé des propositions $A$, $I$, $I$ :

(1.5) Toutes les bêtes venimeuses sont dangereuses
Quelques serpents sont des bêtes venimeuses
donc quelques serpents sont dangereux.

Il a la forme :

(1.6) Tous les $M$ sont $P$
Quelques $S$ sont $M$
donc quelques $S$ sont $P$,

et, plus succinctement :

(1.7) *MaP*
*SiM*
*SiP.*

On le représente par le diagramme 1.11.

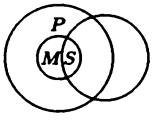



**Figure 1.11** *Syllogisme 1.7*

Les diagrammes d'Euler présentent un inconvénient : il arrive qu'un même syllogisme admette plusieurs représentations géométriques différentes mais équivalentes. R. Blanché donne l'exemple suivant (Blanché, 1970, p. 237) :

(1.8) Aucun $M$ n'est $P$
Quelque $S$ est $M$
donc Quelque $S$ n'est pas $P$.

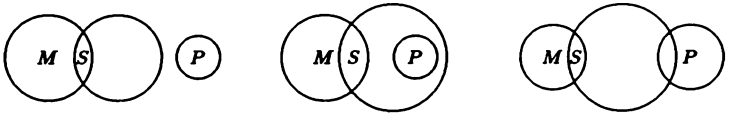



**Figure 1.12** *Syllogisme 1.8*

La particulière "Quelque $S$ n'est pas $P$" comporte une certaine indétermination : on ne précise pas si $P$ est totalement exclu de $S$, totalement inclus dans $S$ ou partiellement inclus dans $S$. La figuration géométrique tranche et impose arbitrairement une des trois figurations qui correspondent respectivement à ces trois lectures de "Quelque $S$ n'est pas $P$".

### 1.3.5. *Forme et contenu*

Pour obtenir la forme d'un syllogisme, on remplace par des marque-place les termes occupant la position du sujet et du prédicat dans les prémisses et la conclusion, en prenant soin de remplacer deux occurrences d'un même terme par deux occurrences d'un même marque-place et deux occurrences de termes différents par deux occurrences de marque-place différents. Remarquons que l'on ne peut substituer à des marque-place que des expressions appartenant à la même catégorie syntaxique que le marque-place. Dans le cas des syllogismes étudiés jusqu'ici, les marque-place sont des marque-place de termes généraux. La classe de substitution des termes généraux ne se limite cependant pas aux noms communs. Elle s'étend aussi aux adjectifs ("humain"), aux verbes intransitifs ("marche"), c'est-à-dire à tout ce qui peut être mis à la place de S, M, P sans produire d'expression syntaxiquement mal formée. On exclut donc les conjonctions, les prépositions ou les phrases complètes, mais non les périphrases comme "points qui sont situés sur le contour d'un cercle ou sur la surface d'une sphère". Si un syllogisme est valide, tous les syllogismes de même forme le sont aussi. Il en va de même pour les syllogismes non valides. Ce trait distingue les syllogismes qui appartiennent, de ce fait, à la logique formelle, de raisonnements comme l'inférence 1.9 :

(1.9) Quelques tableaux sont verts
donc quelques tableaux sont colorés.

Ce dernier raisonnement n'est pas *formellement* valide.

Les catégories syntaxiques dont on a besoin pour faire la théorie du syllogisme diffèrent de celles de la grammaire traditionnelle. Nous venons de voir que le théoricien du syllogisme ne distingue pas les noms communs des verbes intransitifs et des adjectifs. En revanche, il distingue les termes généraux qui s'appliquent à des individus comme "cher", "bon marché", de ceux qui s'appliquent à des classes comme "nombreux", "rare".

Le sophisme

(1.10) Tout ce qui est rare est cher
Les van Gogh bon marché sont rares
donc les van Gogh bon marché sont chers

repose sur la méconnaissance de cette distinction, méconnaissance qui conduit à un usage équivoque du verbe-copule "sont".

### 1.3.6. *Validité et vérité*

Un syllogisme est valide si et seulement si la vérité de la conclusion suit nécessairement de la vérité des prémisses. Les syllogismes valides transmettent la vérité des prémisses aux conclusions. Cette définition exclut qu'un syllogisme valide puisse avoir ses deux prémisses vraies et sa conclusion fausse. Elle n'exclut pas, en revanche, qu'un raisonnement valide ait une conclusion fausse. Mais dans ce cas, au moins une de ses prémisses sera fausse. Le syllogisme 1.11 est valide et a cependant une conclusion fausse :

(1.11) Tous les animaux marins sont des poissons
Toutes les baleines sont des animaux marins
donc toutes les baleines sont des poissons.

### 1.3.7. *Validité et évidence*

Un syllogisme valide, et d'ailleurs tout raisonnement valide, *transmet* la vérité des prémisses aux conclusions. Transmet-il autre chose ? L'un des rôles du raisonnement, c'est de prouver une proposition qui n'est pas évidente à partir de propositions qui sont évidentes. Pour jouer ce rôle de preuve ou de démonstration, le raisonnement doit faire plus que de transmettre la vérité des prémisses aux conclusions, il doit aussi transmettre l'évidence. Il faut distinguer soigneusement la transmission de la vérité de la transmission de l'évidence. Un raisonnement qui n'est pas probant, ses prémisses n'étant pas plus évidentes que sa conclusion, peut néanmoins être valide. Considérons le raisonnement 1.12 :

(1.12) Tous les arbres sont des plantes
donc tous les arbres sont des plantes.

Ce raisonnement a la forme

(1.13) Tous les A sont B
donc tous les A sont B.

Tous les raisonnements de cette forme transmettent la vérité de la prémisse à la conclusion, mais aucun d'eux ne peut servir de preuve. Les raisonnements de ce type sont donc *valides* sans être *probants*. Un raisonnement dans lequel la conclusion figure parmi les prémisses est vicieux dans la mesure où il est présenté comme *preuve* de sa conclusion. Cette faute, qui est une faute argumentative, mais non une faute de logique au sens strict, s'appelle *pétition de principe* ou *cercle vicieux*.

Stuart Mill soutient que tout syllogisme contient un cercle vicieux :

> Il doit, d'abord, être accordé que dans tout syllogisme, considéré comme argument prouvant une conclusion, il y a une *petitio principii*. Quand on dit :

Tous les hommes sont mortels
Socrate est homme
donc Socrate est mortel,

les adversaires de la théorie du syllogisme objectent irréfutablement que la proposition "Socrate est mortel" est présupposée dans l'assertion plus générale "Tous les hommes sont mortels", que nous ne pouvons pas être assurés de la mortalité de tous les hommes à moins d'être déjà certains de la mortalité de chaque homme individuel (Stuart Mill, 1843, tr. fr. 1866, p. 205).

### 1.3.8. *Le raisonnement déductif est-il stérile ?*

Cette critique du syllogisme qui remonte à l'Antiquité — elle est déjà présente chez Sextus Empiricus au III$^{e}$ siècle de notre ère — semble atteindre toutes les inférences déductives et pas seulement le syllogisme car dans tout raisonnement déductif valide, la conclusion est contenue dans les prémisses. On a tenté d'y répondre en disant, comme le fit l'archevêque Whately, contemporain de Stuart Mill, que la conclusion d'un syllogisme n'est contenue qu'*implicitement* dans les prémisses et que le raisonnement est fécond et informatif dans la mesure où il déploie et développe les assertions enveloppées dans les prémisses et nous fait reconnaître la portée de ce que nous affirmons dans les prémisses. Stuart Mill juge cette réponse inacceptable :

> Quand vous admettez la prémisse majeure, vous affirmez la conclusion; mais, dit l'archevêque Whately, vous ne l'affirmez qu'implicitement; ce qui veut dire, sans doute, qu'on l'énonce sans en avoir conscience, sans le savoir. Mais, s'il en est ainsi, la difficulté revient sous une autre forme. Ne devriez-vous pas la connaître ? Quel droit avez-vous d'affirmer la proposition générale sans vous être assuré de la vérité de ce qu'elle contient ? (Stuart Mill, 1943, tr. fr. 1866, 1988, p. 207)

L'objection de Stuart Mill n'est dirimante (Eaton, 1931, p. 148) que si nous interprétons la proposition générale "Tous les hommes sont mortels" comme synonyme d'une conjonction de propositions singulières "Thalès est mortel et Parménide est mortel et Socrate est mortel etc.", car dans ce cas, la conclusion est littéralement contenue dans les prémisses. Cependant rien ne nous oblige à interpréter ainsi la proposition universelle. A côté de l'interprétation en extension, il y a place pour une interprétation intensionnelle. Selon cette dernière interprétation, affirmer que tous les hommes sont mortels, c'est affirmer que la propriété de mortalité est incluse dans celle d'humanité. Or cette inclusion de propriété peut être connue sans passer tous les hommes en revue : il suffit de constater qu'ils sont formés de matière organique corruptible.

Lu en intension, le syllogisme cité par Mill devient :

(1.14) La propriété de mortalité appartient à celle d'humanité
La propriété d'humanité appartient à Socrate
donc la propriété de mortalité appartient à Socrate.

L'objection selon laquelle le raisonnement déductif, en général, est stérile parce qu'il se borne à exprimer de manière *explicite* dans la conclusion ce qui était *implicite* dans les prémisses peut être réfutée sans faire appel à la distinction entre extension et intension dont nous venons de faire un usage ponctuel pour répondre à la critique particulière de Stuart Mill. Il y a des déductions qui accroissent notre connaissance. Les conclusions obtenues par ces déductions sont bien *contenues* dans les prémisses, et dans les définitions posées au départ, mais, comme l'écrit Frege dans les *Grundlagen der Arithmetik*, "[...] elles le sont comme une plante l'est dans la graine, non pas comme une poutre l'est dans la maison" (Frege, 1884, 1969, p. 212). Frege ajoute :

> Souvent, plusieurs définitions sont nécessaires à la démonstration d'une proposition; elle n'est donc contenue dans aucune d'elles prises séparément, bien qu'elle découle de leur conjonction par le seul fait de la logique pure (*Ibid.*).

Nous verrons plus loin qu'une déduction n'est généralement pas une opération par laquelle on *extrait* ou *dévoile* une conclusion contenue et cachée dans les prémisses, mais une *construction* qui oblige à *transformer* les prémisses et non pas seulement à les *décomposer* en leurs constituants.

## 1.4. La tradition stoïcienne

Considérons les deux raisonnements 1.15 et 1.16.

(1.15) Tous les *hommes* sont *mortels*
donc tous les non *mortels* sont non *hommes*.

(1.16) S'*il fait jour*, *il fait clair*
donc s'*il* ne *fait* pas *clair*, *il* ne *fait* pas *jour*

Ces deux raisonnements ont respectivement la forme

(1.17) Tous les $A$ sont $B$
donc tous les non $B$ sont non $A$

et

(1.18) Si $p$, alors $q$
donc si non $q$ alors non $p$.

Il y a quelque chose de commun à ces deux raisonnements : une *permutation* et un *changement de signe*. Cette double opération porte le nom de *contraposition*. Une différence importante doit être signalée : les *marque-place du schéma* représentent des *termes* dans le premier raisonnement, alors qu'ils représentent des *propositions* dans le second. Les raisonnements qu'étudie Aristote sont des

raisonnements de la première espèce : les *unités* les plus petites sont des *termes* (noms, verbes) et les *symboles logiques* sont des symboles *intrapropositionnels* comme la copule ou la négation interne portant sur un terme ou sur la copule.

Les Stoïciens inaugurent l'étude de raisonnement de la deuxième espèce : les *unités* les plus petites sont des *propositions* et les symboles logiques sont des symboles *interpropositionnels* comme les connecteurs binaires ou la négation externe, connecteur unaire qui porte sur une proposition.

Les schémas d'inférence de la logique aristotélicienne reposaient sur quatre constantes logiques : deux copules, "est" et "n'est pas" qui diffèrent par la *qualité*, et deux marqueurs de *quantité* ("tout" et "aucun", signe de la quantité universelle, et "quelque", signe de la quantité particulière). Les schémas d'inférence de la logique stoïcienne reposent sur d'autres constantes logiques : les connecteurs "si... alors...", "ou", "et", "non".

### 1.4.1. *Le conditionnel vérifonctionnel*

Les Stoïciens (cf. Vuillemin 1989) ont dégagé deux interprétations du connecteur "Si... alors...". Selon l'un de ces sens (proposé par Philon de Mégare), une proposition conditionnelle de la forme "Si $p$ alors $q$", où "$p$" et "$q$" sont des marque-place de phrases déclaratives, est vraie dans trois cas :

- lorsque l'antécédent et le conséquent sont vrais,
- lorsque l'antécédent et le conséquent sont faux,
- lorsque l'antécédent est faux et le conséquent vrai.

Elle est fausse dans un cas, lorsque l'antécédent est vrai et le conséquent faux.

Selon cette acception, la vérité ou la fausseté du conditionnel dépend uniquement de la vérité ou de la fausseté de l'antécédent et du conséquent. Elle ne *dépend* pas de l'existence d'une connexion quelconque entre l'antécédent et le conséquent. Appelons le vrai et le faux des "valeurs de vérité". Nous pouvons exprimer cette dépendance en termes mathématiques et dire que la valeur de vérité de la conditionnelle est *fonction* de la valeur de vérité de l'antécédent et du conséquent. Le conditionnel philonien est donc, avant la lettre, un foncteur de vérité, c'est-à-dire un connecteur vérifonctionnel que l'on interprète en lui associant une fonction de vérité.

Tout esprit non prévenu à qui on demanderait d'assigner une valeur de vérité au conditionnel "Si Paris est en Angleterre, $2 + 2 = 4$" répondrait que la question est absurde et que ce conditionnel n'a pas de valeur de vérité du tout. Philon réagit autrement parce qu'il met à la base de sa logique le *principe de bivalence* (Toute proposition est vraie ou fausse). Il admet implicitement que tout ce qui n'est pas faux est vrai. Or la proposition "Si Paris est en Angleterre, alors $2 + 2 = 4$" n'est pas fausse donc elle est vraie. Au fond, Philon prend la conjonction "Si... alors" dans un sens affaibli qui est bien exprimé par la

définition suivante :

$$\text{Si } p \text{ alors } q \quad =_{def} \quad \text{Non } (p \text{ et non } q).$$

Le signe $=_{def}$ se lit "est défini par"[2]. En d'autres termes, "il est vrai que si $p$ alors $q$" équivaut à "il est faux que $p$ soit vrai et $q$ faux ensemble".

Face à cette acception, il y en a d'autres. Diodore impose des conditions plus fortes à la vérité d'une proposition conditionnelle. Selon lui, pour qu'une telle proposition soit vraie, il ne suffit pas que la vérité de l'antécédent ne coexiste pas avec la fausseté du conséquent; il faut, par surcroît, qu'une telle coexistence soit *impossible*. En d'autres termes, il faut que la vérité de l'antécédent soit incompatible avec la fausseté du conséquent. Il reste à expliquer en quoi consiste cette incompatibilité.

### 1.4.2. *Les cinq schémas de base des Stoïciens*

Le syllogisme catégorique d'Aristote comportait deux prémisses et une conclusion. Il consistait dans une opération quasi algébrique de simplification : l'élimination du moyen terme. Les cinq schémas des Stoïciens sont également des transformations d'effacement, mais ici, ce qui est éliminé, ce n'est plus un *terme* mais une *proposition* et un *connecteur*. On peut garder l'expression de *majeure* pour désigner la prémisse contenant les deux propositions et l'expression de *mineure* pour désigner la prémisse contenant la proposition qui sera éliminée.

[2]Le symbole $\equiv$ est une conjonction grammaticale (un connecteur) qui unit des phrases. Le symbole $=_{def}$ est un prédicat qui unit des noms de phrases. Généralement, on forme le nom d'une phrase $p$ à partir de celle-ci en la plaçant entre guillemets ("$p$"). Pour alléger l'expression, nous supprimons systématiquement les guillements quand le contexte indique clairement que le signe (phrase ou mot) est *désigné* plutôt qu'*utilisé*. Les phrases à gauche et à droite du signe $=_{def}$ se nomment elles-mêmes (convention d'*autonymie*). En pratique, nous conserverons les guillemets uniquement pour les expressions de la langue naturelle ou semi-naturelle (contenant des parenthèses et des marque-place) qui ne sont pas *utilisées* mais *désignées*.

| | *Schémas d'inférence* | *Exemples* |
|---|---|---|
| (1.19) | **Modus ponens** | |
| | Si $p$ alors $q$ | Si la sueur traverse la peau, il y a des pores |
| | Or $p$ | Or la sueur traverse la peau |
| | Donc $q$ | Donc il y a des pores |
| (1.20) | **Modus tollens** | |
| | Si $p$ alors $q$ | Si la grippe est présente, la fièvre est présente |
| | Or non $q$ | Or la fièvre n'est pas présente |
| | Donc non $p$ | Donc la grippe n'est pas présente |
| (1.21) | **Forme 1 du syllogisme disjonctif** | |
| | pas simultanément $p$ et $q$ | Il n'est pas simultanément le cas que la fièvre soit présente et que l'organisme soit en bonne santé |
| | Or $p$ | Or la fièvre est présente |
| | Donc pas $q$ | Donc l'organisme n'est pas en bonne santé |
| (1.22) | **Forme 2 du syllogisme disjonctif** | |
| | $p$ ou $q$ (mais pas les deux) | Ou deux est pair, ou deux est impair |
| | Or $p$ | Or deux est pair |
| | Donc non $q$ | donc deux n'est pas impair |
| (1.23) | **Forme 3 du syllogisme disjonctif** | |
| | $p$ ou $q$ (mais pas les deux) | Ou deux est pair, ou deux est impair |
| | Or non $q$ | Or deux n'est pas impair |
| | Donc $p$ | Donc deux est pair |

*Remarque.* Les schémas 1.22 et 1.23 éliminent le connecteur "ou". Le "ou" dont il s'agit, c'est le "ou" exclusif — *aut* en latin — qui figure dans certains menus de restaurant. Si on remplaçait le "ou" exclusif par un "ou" inclusif — *vel* en latin —, le dernier schéma continuerait à être valide, mais pas l'avant-dernier.

### 1.4.3. *Variables et marque-place*

Les lettres $p$ et $q$ ne sont plus seulement des marque-place destinés à dégager la forme logique en expulsant le contenu extra-logique. Ce sont d'authentiques variables auxquelles on peut appliquer la règle de substitution pour engendrer

de nouveaux schémas d'inférence valides à partir de ceux qu'on a déjà. Il faut distinguer soigneusement deux sortes de substitutions : jusqu'ici nous avons appliqué la substitution pour dégager la forme logique d'un raisonnement — cette sorte de substitution exige que l'on remplace les occurrences de la même proposition par des occurrences du même marque-place et que l'on remplace des occurrences d'une proposition différente par des marque-place différents. A présent, nous rencontrons une nouvelle sorte de substitution : la substitution comme règle d'inférence. Au lieu de substituer des marque-place à des propositions de la langue naturelle, on substitue des marque-place à d'autres marque-place (par exemple $p$ à $q$) ou des schémas (marque-place combinés à des marque-place par des connecteurs) à d'autres marque-place (par exemple "$p$ et $q$" à $p$). Il s'agit d'une règle d'inférence au deuxième degré qui permet d'engendrer des schémas dérivés à partir des cinq schémas primitifs. Cette règle transporte la validité des schémas de départ aux schémas d'arrivée à condition que soit respectée une contrainte d'uniformité. Cette nouvelle espèce de substitution obéit à des contraintes propres :

- on ne peut appliquer la substitution qu'au sein d'un schéma primitif ou d'un schéma obtenu à partir des schémas primitifs,
- on peut remplacer des marque-place différents (par exemple $p$ et $q$) par le même substitut.

(1.24)

| | | |
|---|---|---|
| Prémisse I | Si ($p$ et $q$) alors $r$ | |
| Partie de prémisse II | non $r$ | schéma 2 |
| Conclusion provisoire | non ($p$ et $q$) | |
| Partie de prémisse II | $p$ | schéma 3 |
| Conclusion | Donc non $q$ | |

Cette démonstration qui remonte aux Stoïciens préfigure les démonstrations obtenues à l'aide de la déduction naturelle qui sera exposée dans le chapitre 3.

### 1.4.4. *Forme inférentielle et forme propositionnelle*

Les Stoïciens distinguent les inférences telles que, par exemple,

(1.25) S'il fait jour, il fait clair
Or il fait jour
Donc il fait clair,

de la proposition conditionnelle que l'on obtient quand on remplace le "donc" par un "si ... alors..." :

(1.26) Si (s'il fait jour il fait clair) et il fait jour, alors il fait clair.

Cette distinction est importante : un raisonnement est une opération de l'esprit par laquelle on passe d'une ou plusieurs prémisses à la conclusion. Il ne peut être dit vrai ou faux car ce passage d'une proposition à une autre n'est pas lui-même une proposition. Cependant, comme l'ont observé les Stoïciens, il y

a une proposition conditionnelle qui correspond à toute inférence (la proposition conditionnelle qu'on obtient en prenant les prémisses de l'inférence comme antécédent et la conclusion comme conséquent). On peut rattacher la ***validité*** d'une inférence à la ***vérité*** de la proposition conditionnelle qui lui correspond et de même la ***validité*** d'une forme d'inférence à la ***validité*** de la forme propositionnelle (proposition contenant des variables propositionnelles) correspondante. Le schéma du ***modus ponens*** lui-même, à savoir

(1.27) Si $p$ alors $q$
Or $p$
Donc $q$,

est valide s'il est ***impossible*** que l'antécédent $A$ de la forme propositionnelle conditionnelle soit vrai et que le conséquent $B$ soit faux. En d'autres termes, 1.27 est valide si la forme propositionnelle 1.28 est valide (vraie pour toutes les interprétations de $p$ et $q$).

(1.28) $\underbrace{\text{Si } p \text{ et que si } p \text{ alors } q}_{A}, \underbrace{\text{alors } q}_{B}$

Le "si... alors..." principal reçoit donc l'interprétation diodoréenne plutôt que philonienne.

Si nous mettons sous forme propositionnelle le schéma d'inférence dérivé 1.24, nous obtenons 1.29 :

(1.29) Si (si ($p$ et $q$) alors $r$) et si non $r$ et $p$, alors non $q$.

### 1.4.5. *Proposition conditionnelle et implication*

Au lieu de dire "Si $p$ alors $q$, or $p$ donc $q$", on dit parfois "$p$ implique $q$, or $p$ donc $q$". Ces deux constructions ne sont pas interchangeables. Elles appartiennent, en effet, à des couches différentes du langage qu'il importe de ne pas confondre.

Considérons les deux phrases suivantes :

(1.30) Socrate est chauve,

(1.31) "Socrate" a sept lettres.

La première parle de l'individu Socrate. La deuxième parle de son nom. Nous avons des noms propres pour désigner des choses et des noms propres pour désigner les noms propres (le nom propre entre guillemets est une sorte de nom propre du nom). De même nous avons des phrases qui parlent du monde, comme la phrase

(1.32) la terre est ronde,

et des phrases qui parlent des phrases comme la phrase

(1.33) "la terre est ronde" est vraie.

Enfin, on peut distinguer le langage-objet, qui sera l'objet de notre étude, du métalangage qui est un deuxième langage dans lequel on parle du premier. La phrase

(1.34) S'il fait clair, alors il fait clair

appartient au langage-objet tandis que la phrase

(1.35) "S'il fait clair, alors il fait clair" est une vérité logique

appartient au métalangage.

La phrase

(1.36) "Il fait clair" implique logiquement "Il fait clair"

appartient au métalangage également.

La phrase 1.35 a d'ailleurs été proposée comme une définition de la phrase 1.36. On voit à présent que l'implication diffère du conditionnel : affirmer qu'une proposition en implique une autre, c'est faire plus que d'exprimer un énoncé de forme conditionnelle, c'est dire, en outre, de ce conditionnel qu'il est une *vérité de logique*, notion qui sera expliquée dans la section suivante.

### 1.4.6. *Les vérités de logique*

Tant qu'on n'a pas défini la notion de *vérité de logique*, on ne voit guère l'utilité de la reconnaissance par les Stoïciens d'un lien entre la *validité* des *inférences* et la *vérité* des *propositions* conditionnelles correspondantes (les propositions conditionnelles obtenues en remplaçant la particule "donc" par la conjonction "si ... alors..."). La situation changera quand on disposera d'une définition parfaitement claire de la notion de vérité de logique pour la logique des propositions. Dès à présent il est possible de la formuler : une proposition est une vérité de logique de la logique des propositions limitée aux connecteurs vérifonctionnels si et seulement si elle est vraie quelle que soit la valeur de vérité des propositions atomiques qui la composent. On appelle proposition atomique une proposition qui ne contient pas de connecteurs tels que "non", "et", "ou (inclusif)", "ou (exclusif)", "si ... alors ...", "si et seulement si", etc.[3] Par exemple, "Il fait clair", "Socrate est chauve". On appelle *tautologies* les vérités de logique du calcul des propositions.

Considérons les deux énoncés suivants :

(1.37) S'il fait jour alors s'il fait jour, il fait jour,

(1.38) Si (S'il fait jour, il fait jour), alors il fait jour.

La proposition 1.37 est une vérité de logique qui a la forme "Si $p$ alors (si $p$ alors $p$)". La proposition élémentaire "il fait jour" est vraie ou fausse.

[3] Il y quatre connecteurs vérifonctionnels unaires en logique bivalente et seize connecteurs vérifonctionnels binaires.

Si elle est vraie, la proposition composée "s'il fait jour, il fait jour" l'est aussi puisqu'un conditionnel dont le conséquent est vrai est vrai. La proposition 1.37 l'est donc aussi, pour la même raison. Si la proposition élémentaire "il fait jour" est fausse, la proposition 1.37 est vraie parce que tout conditionnel dont l'antécédent est faux est vrai.

La proposition 1.38, au contraire, n'est pas une vérité de logique. En effet, elle n'est pas vraie quelle que soit la valeur de vérité de ses atomes : elle est fausse si "il fait jour" est faux car dans ce cas elle a un antécédent vrai et un conséquent faux.

### 1.4.7. *Logique aristotélicienne et logique stoïcienne*

La contribution centrale d'Aristote et de ses commentateurs et continuateurs à la logique est sa théorie du syllogisme. Cette théorie permet d'étudier exhaustivement les 256 formes possibles du syllogisme catégorique, de séparer celles qui sont valides de celles qui ne le sont pas et de justifier les syllogismes valides.

Malgré son importance indiscutable, la logique aristotélicienne reste limitée : elle n'est qu'une *partie* de la logique des termes. Elle ne permet pas de décider, par exemple, si le raisonnement 1.39 qui appartient pourtant à la logique des termes est valide.

(1.39) Si tous les visiteurs suspects sont coupables, alors quelques visiteurs ne sont pas suspects
Or tous les visiteurs sont suspects ou tous les visiteurs sont coupables
Donc si tous les visiteurs coupables sont suspects, alors quelques visiteurs qui ne sont pas coupables sont suspects.

La logique stoïcienne, par le biais de la règle de substitution, s'étend, au contraire, à un nombre *infini* de schémas d'inférence propositionnels.

Les logiciens contemporains demandent davantage; ils cherchent à établir que leur logique est *complète*, c'est-à-dire capable de reconnaître tous les schémas valides d'inférence formulables dans cette logique. O. Becker a montré que la logique stoïcienne avait cette propriété pour ce qui est des schémas formés à l'aide de certains connecteurs.

La logique des propositions est plus fondamentale que la logique des termes, en ce sens que la logique des termes la présuppose. En effet, on a besoin des schémas d'inférence de la logique des propositions pour démontrer, notamment, que certains syllogismes qui ne sont pas des instances de la première figure sont valides, ce qu'on fait en les réduisant à des syllogismes de la première figure. Considérons, par exemple, le syllogisme 1.40 :

(1.40) Tous les poissons ont des branchies
Quelques animaux marins n'ont pas de branchies
donc quelques animaux marins ne sont pas des poissons.

Pour faciliter la réduction, nous représenterons "poisson" qui est le sujet de la conclusion par $M$, "a des branchies" qui est le terme majeur par $P$, dérogeant ainsi à notre convention. Nous obtenons :

(1.41) $MaP$
$SoP$
donc $SoM$.

En vertu des inférences immédiates inventoriées dans le carré logique, "Quelques animaux marins n'ont pas de branchies" ($SoP$) équivaut à "Il n'est pas le cas que tous les animaux marins aient des branchies" qui a la forme "Non $SaP$". De même la conclusion de forme $SoM$ est équivalente, pour la même raison, à "Non $SaM$". Moyennant les transformations autorisées par ces équivalences, le syllogisme en question peut être réécrit sous la forme suivante :

(1.42) $MaP$
Non $SaP$
donc non $SaM$.

Or le syllogisme ainsi réécrit découle du syllogisme "parfait" de la première figure ($Map$, $SaM$ donc $SaP$) en vertu du principe suivant que les Stoïciens dérivaient de leurs cinq schémas fondamentaux. Dans le schéma 1.43,

(1.43) Que [si ($p$ et $q$) alors $r$]
implique que [si ($p$ et non $r$) alors non $q$]

il suffit de faire les substitutions suivantes :

$MaP$ / $p$
$SaM$ / $q$
$SaP$ / $r$.

En d'autres termes :

(1.44) (Si ($MaP$ et $SaM$) alors $SaP$)
implique
(Si ($MaP$ et $\underbrace{\text{non } SaP}_{SoP}$) alors $\underbrace{\text{non } SaM}_{SoM}$).

## 1.5. L'apport de Leibniz

### 1.5.1. *De la logique formelle à la logique formalisée*

La logique d'Aristote et celle des Stoïciens sont formelles. La présence de marque-place et de variables dans les schémas de raisonnement l'atteste, mais ces marque-place et ces variables sont combinés avec des expressions de la

langue naturelle qui conservent leur signification (les copules "Tous...sont...", "Quelque...n'est pas...", les conjonctions "...ou...", "si...alors..."). Tant qu'on raisonne avec des mots chargés de sens, des désaccords peuvent surgir sur le sens que nous leur donnons. Il en irait tout autrement si, dans un raisonnement, on ne devait prendre en compte que les signes, ou caractères, abstraction faite de leur sens. La seule question qui continue à se poser, c'est de savoir reconnaître les caractères, de savoir dire quand nous nous trouvons devant deux *occurrences* ou *échantillons* du même *type* de signe ou devant deux occurrences ou échantillons de types différents (l'opposition *type-occurrence* est due au logicien américain Ch. S. Peirce). En revanche, nous avons écarté toute possibilité d'équivoque ou de malentendu puisque les mots n'ont plus de signification et qu'ils ne renvoient plus à des référents.

Leibniz a introduit l'idée du raisonnement conçu comme un calcul et celle d'une logique qui n'est pas seulement formelle mais *formalisée* :

> Je crois qu'on ne pourra jamais mettre fin aux controverses [...] si l'on ne revient pas des raisonnements compliqués à des *calculs* simples, et des mots ayant des significations vagues et incertaines à des *caractères* bien déterminés. Car il arrivera alors que tout paralogisme ne soit rien de plus qu'une *erreur de calcul* [...]. Dès lors, quand surgiront des controverses, inutile d'instituer une discussion entre deux philosophes, pas plus qu'on ne le fait entre deux calculateurs. Car il suffira de prendre la plume à la main ou de s'asseoir devant les abaques et, après avoir au besoin convoqué un ami, de se dire l'un à l'autre : "calculons !" (Leibniz, ed. 1890, p. 200, cité par Blanché, 1970, p. 214).

Inventeur du projet d'un *calculus ratiocinator*, Leibniz est également l'auteur du projet d'une *lingua characteristica universalis*, c'est-à-dire d'une langue artificielle, dotée de signification comme la langue naturelle, mais pourvue de toutes les bonnes propriétés que nous voudrons lui donner comme celle dont Raymond Lulle avait conçu l'idée. Les deux projets, restés à l'état d'ébauche chez Leibniz, seront réalisés ultérieurement.

### 1.5.2. *La logique de l'identité*

Un des apports décisifs de Leibniz à la logique est la formulation des lois de la logique de l'identité. Par logique de l'identité, nous entendrons les lois relatives aux raisonnements dont la validité ou la non-validité dépend du terme "=". Leibniz a énoncé la règle suivante qu'il exprime en latin et dont l'interprétation exige le plus grand soin : *Eadem sunt, quae mutuo substitui possunt, salva veritate*, ce qui, traduit littéralement, signifie "Sont les mêmes, ceux qui peuvent être substitués l'un à l'autre, la vérité restant sauve". Une question se pose immédiatement au lecteur de la loi de Leibniz : le terme général "être le même que" s'applique-t-il aux mêmes êtres que le terme général "peut être substitué

à" ? A première vue, les êtres qui peuvent être identiques ou non, ce sont des objets physiques, des objets mathématiques, voire des mots, mais *seuls* des mots peuvent être substitués l'un à l'autre dans une phrase. Il ne faut pas confondre les termes qui peuvent être placés à gauche et à droite du signe "=" avec les objets que ces termes dénotent. Cette confusion a été commise par l'un des plus grands logiciens. Russell, en effet, traduit ainsi la formule latine de Leibniz :

> Si *a* est identique à *b*, tout ce qui est vrai de l'un est vrai de l'autre et ils peuvent être substitués l'un à l'autre dans n'importe quelle proposition, sans altérer la vérité ou la fausseté de cette proposition (Russell, 1905, 1956, p. 47).

Dans l'antécédent de cet énoncé conditionnel, Russell fait référence aux *objets* *a* et *b* et, subrepticement, dans le conséquent, il fait référence aux *termes* qui les désignent. Ce sont en effet les *termes* et non les *objets* qui sont inter-substituables au sein de la proposition. La bonne formulation est la suivante : "si deux objets sont identiques, les termes qui les désignent sont interchangeables, la vérité restant sauve".

De ce que le maître de Platon (Socrate) est identique au mari de Xanthippe, il découle en vertu de la loi de l'identité formulée par Leibniz que la proposition vraie

(1.45) Le maître de Platon a bu la ciguë

reste vraie quand on substitue au terme singulier "Le maître de Platon", le terme singulier "Le mari de Xanthippe".

### 1.5.3. *La relation d'identité*

L'identité est-elle une relation entre objets ou entre noms d'objets ? Frege a posé cette importante question qui l'a conduit à faire des distinctions qui permettent de préciser les limites de validité de la loi de Leibniz, limites que Leibniz connaissait, mais qu'il n'a pas exposées aussi clairement que Frege (H. Ishiguro, 1972). Frege a commencé par soutenir que l'identité est une relation entre objets et que, dès lors, un énoncé de la forme "$a = b$" est vrai si et seulement si l'objet dénoté par "$a$" est le même objet que l'objet dénoté par "$b$".

Cette vue n'est pas fausse, mais elle est trop sommaire. Elle ne permet pas d'expliquer la différence entre les jugements d'identité vrais *a priori* et non informatifs comme "le maître de Platon = le maître de Platon" et les jugements d'identité qui ne sont pas vrais *a priori*, mais qui sont informatifs comme "le maître de Platon = le mari de Xanthippe". Frege a réussi à rendre compte de la différence entre les deux jugements d'identité précités au moyen de la distinction entre le *sens* d'une expression et sa *référence*. Les deux termes singuliers "le maître de Platon" et "le mari de Xanthippe" sont coréférentiels —

ils renvoient au même individu —, mais ils ne sont pas synonymes. Les deux jugements d'identité ont même valeur de vérité, ils sont tous les deux vrais, mais ils n'ont pas le même sens, ils expriment des *états des choses différents* et des *propositions* différentes. Dans cet usage, le mot proposition n'est plus synonyme de phrase, ce n'est plus un être linguistique.

La loi de Leibniz qui énonce que les termes coréférentiels sont interchangeables, la vérité restant sauve, est prise en défaut dans certains cas, ce que Leibniz avait observé : l'exemple 1.46 de non-application de la loi est dû à Russell (Russell, 1905).

(1.46) Jocaste = la mère d'Oedipe
Oedipe désirait épouser Jocaste
Oedipe désirait épouser sa mère

L'inférence 1.47 est un autre exemple, dû à Quine (Quine, 1953).

(1.47) Giorgione était appelé ainsi en raison de sa petite taille
Giorgione = Barbarelli
Donc, Barbarelli était appelé ainsi en raison de sa petite taille

La loi de Leibniz ne s'applique pas, au moins sans modification, aux occurrences de termes singuliers qui se produisent dans des *contextes d'attitude propositionnelle* (attitudes exprimées par les verbes "croire", "savoir", "désirer", etc.) ni dans les *contextes de citation*. Dans l'exemple de Quine, on se trouve en présence d'un contexte de citation masqué. En effet, "Giorgione était appelé *ainsi*" est en fait mis pour "Giorgione était appelé "Giorgione" ".

La loi de Leibniz s'applique sans poser de problèmes à la logique des prédicats (voir chapitre 5), laquelle suffit à la démonstration en mathématique. Il n'est pas nécessaire, à ce stade, de préciser davantage le champ dans lequel la loi de Leibniz est valide.

Même les exemples de Russell et de Quine, cités plus haut, ne sont pas des contre-exemples indiscutables à la loi de Leibniz. La conclusion est valide dans une certaine lecture de "sa mère" et de "ainsi". On peut dire que "sa mère" est un terme singulier à mettre dans la bouche du locuteur et pas d'Oedipe et soutenir que *ainsi* renvoie au diminutif "Giorgione".

## 1.6. La logique de Boole

Si nous comparons le terme général "homme" au terme général "non-homme" en nous concentrant sur l'intension, une différence nous frappe immédiatement. Le premier connote un ensemble de propriétés. Le second est dépourvu d'intension assignable pour un entendement humain. "Ce serait pour lui une tâche impossible que d'abstraire les caractères communs de tout ce qui n'est pas homme — par exemple : du triangle, de la mélancolie et de l'acide sulfurique — pour les combiner dans le concept *non-homme*" (Schröder, cité par Cl. Imbert, 1979, pp. 641–642). On serait bien en peine de lui donner

une définition. On comprend que, se plaçant au point de vue de l'intension, Aristote ait considéré les termes négatifs comme des termes *indéterminés* et qu'il les ait rejetés purement et simplement.

G. Boole (1815–1864) a montré qu'il est légitime, en revanche, de mettre sur le même pied les termes généraux positifs et négatifs traités en *extension*. Pour justifier ce traitement égal, il a introduit la notion de *classe complémentaire* ainsi que les notions de *classe universelle* et de *classe vide*.

Le contraste entre termes positifs et termes négatifs au plan de l'*intension* qui a été *justement souligné* par Schröder doit cependant être assorti de certaines réserves. Il existe des termes généraux dont l'intension ne connote pas un ensemble de propriétés que possèderaient *tous* les individus auxquels ces termes s'appliquent. Wittgenstein donne l'exemple du mot "jeu". Tous les jeux ont quelques propriétés en commun, mais aucune propriété n'est commune à tous les jeux. Il n'y a entre les jeux qu'une "ressemblance de famille", ressemblance qui n'est pas une relation transitive.

Le théorème de Watanabe vient conforter et généraliser l'idée de Wittgenstein :

> Quand on se sert d'un concept, on entend d'habitude qu'il y a un groupe d'objets tels que deux membres quelconques du groupe se ressemblent plus qu'un membre et un non-membre ne se ressemblent. Deux moineaux se ressemblent beaucoup, tandis qu'un moineau et une rose ne se ressemblent pas. Or, il serait naturel que le prédicat "se ressembler" puisse se traduire par "partager beaucoup de prédicats [c'est-à-dire de termes généraux] en commun". Mais cette interprétation se trouve être fausse [...] Le théorème énonce que le nombre de prédicats qui s'appliquent à deux objets non identiques est un nombre fixe indépendant de ces deux objets (Watanabe, 1965, p. 48).

Revenons à Boole. La contribution majeure de Boole à la logique peut se résumer ainsi. Il a montré :

- que la logique peut prendre la forme d'un calcul dont la correction "ne dépend pas de l'interprétation des symboles employés mais seulement des lois qui régissent la manière dont ces symboles sont combinés",
- que l'algèbre appliquée jusque-là aux nombres pouvait être étendue aux classes,
- que les quatre propositions fondamentales d'Aristote (interprétées en extension) pouvaient être exprimées au moyen d'équations de l'algèbre ainsi étendue,
- que la validité ou la non-validité des syllogismes pouvait être établie par un calcul algébrique sur ces équations.

Pour convertir l'algèbre ordinaire qui s'applique aux nombres en une algèbre des classes qui permettra de mathématiser la syllogistique aristotélicienne, Boole adopte les interprétations suivantes (Kilmister, 1967, p. 20) :

- $x \times y$ dénote l'ensemble des membres de $x$ qui sont aussi membres de $y$,
- si $x$ et $y$ n'ont pas de membres en commun, $x + y$ dénote l'ensemble des objets qui appartiennent soit à $x$ soit à $y$,
- si $x$ et $y$ ont des membres en commun, le symbole $x + y$ est admis dans l'exécution des calculs, mais comme il est ce que Boole appelle une "forme non interprétée", il n'est pas autorisé dans la formule finale,
- $1 - x$ dénote tous les objets qui n'appartiennent pas à la classe $x$,
- $x = 0$ signifie que la classe $x$ n'a pas de membres.

Dans l'interprétation qui vient d'être donnée, on a les lois suivantes :

$$x + y = y + x \tag{1.48}$$

$$x \times y = y \times x \tag{1.49}$$

$$x + (y + z) = (x + y) + z \tag{1.50}$$

$$x \times (y \times z) = (x \times y) \times z \tag{1.51}$$

$$x \times (y + z) = (x \times y) + (x \times z) \tag{1.52}$$

Ces lois ressemblent à celles de l'algèbre ordinaire où les variables $x$, $y$ et $z$ prennent leurs valeurs dans les nombres et les symboles "+", "×" désignent respectivement la somme et le produit arithmétique.

On a également la loi :

$$x^2 = x \times x = x \tag{1.53}$$

qui n'a pas de correspondant dans l'algèbre ordinaire, à moins qu'on ne limite les nombres pris comme valeurs des variables à 0 et 1.

Boole a proposé la mise en équations des quatre propositions fondamentales $A$, $E$, $I$, $O$[4] :

| | | | |
|---|---|---|---|
| Tout $X$ est $Y$ | $x(1-y)$ | $=$ | $0$ |
| Aucun $X$ n'est $Y$ | $x \times y$ | $=$ | $0$ |
| Quelque $X$ est $Y$ | $x \times y$ | $\neq$ | $0$ |
| Quelque $X$ n'est pas $Y$ | $x(1-y)$ | $\neq$ | $0$ |

Boole exprime les propositions particulières (1 et 0) sous une autre forme de manière à avoir des égalités partout. Tous les types de raisonnement de la logique aristotélicienne peuvent être représentés dans cette algèbre; le syllogisme y prend la forme de la réduction de deux équations (correspondant aux deux prémisses) à une seule, suivie par l'élimination du moyen terme. Nous n'exposerons pas ici la méthode de Boole.

[4] Il importe de distinguer "O" désignant la proposition négative particulière, de "0" désignant la classe vide.

### 1.6.1. *L'algèbre de Boole-Schröder*

Le système de Boole comporte un certain nombre de défauts, notamment l'usage d'expressions qui n'ont pas une signification constante. Ces défauts seront éliminés par les successeurs de Boole. C'est la version de l'algèbre de Boole due à Schröder qui prévaut aujourd'hui. Au lieu d'utiliser l'opération exprimée par le signe + qui dénote la réunion de classes disjointes, on utilise une opération $\cup$ qui n'impose plus cette restriction et qui dénote, en fait, la réunion ensembliste. L'opération $\cup$ vérifie certaines lois que ne vérifiait pas l'opération désignée par le signe +, par exemple la réunion ($\cup$) se distribue sur l'intersection ($\cap$) alors que l'addition des classes (+) ne se distribuait pas sur le produit des classes ($\times$). Rien ne change pour ce qui est de la transcription algébrique des quatre propositions de la logique traditionnelle, mais on dispose maintenant de méthodes plus commodes pour vérifier la validité des syllogismes, et on dispose de *lois de dualité* qui permettront de donner un fondement à la loi de variation inverse de l'extension et de l'intension.

L'algèbre de Boole vérifie le *principe de dualité* qui s'énonce ainsi : "Tout théorème qu'on peut déduire des axiomes de l'algèbre de Boole demeure vrai quand les symboles $\vee$ et $\wedge$ ainsi que les éléments neutres 0 et 1 sont échangés dans tout l'énoncé". Ce principe permet, par exemple, d'obtenir $\alpha \cup \overline{\alpha} = 1$ à partir de $\alpha \cap \overline{\alpha} = 0$.

Grize remarque qu'il est possible, grâce à la dualité, d'interpréter l'algèbre de Boole en extension ou en compréhension :

> En effet, tandis qu'on a, par exemple, que l'*intersection* de la classe des rectangles avec celle des losanges détermine la classe des carrés, c'est la *réunion* des propriétés des rectangles à celles des losanges qui fournit les propriétés des carrés. De là découle la fameuse loi de la variation inverse de l'extension et de la compréhension des concepts. Il y a "moins" de carrés que de rectangles puisqu'il existe des rectangles qui n'ont pas leurs quatre côtés égaux, mais les carrés jouissent de "plus" de propriétés que les rectangles (Grize, 1967, p. 272).

Comme le fait remarquer J. P. Desclés, cette dualité trouve aujourd'hui des applications en intelligence artificielle :

> [...] la dualité traduirait bien ce que l'on appelle maintenant "le principe d'héritage des propriétés". Ainsi, selon le principe de dualité, tout objet qui "tombe sous" le concept "homme" hériterait de toutes les propriétés qui se trouvent dans son intension; il "tombe" alors lui aussi sous le concept "animal" et sous le concept "bipède"... (Desclés, 1987, p. 60).

Il ne faut cependant pas surestimer la portée de ce rapprochement :

> [...] la dualité et l'héritage des propriétés ne tiennent que pour les *occurrences typiques* ou prototypes du concept [...] Par exemple, un individu unijambiste appartient de plein droit à l'extension de "homme", il tombe

donc sous le concept "homme", mais il *n'hérite pas* pour autant *de toutes* les propriétés qui sont dans l'intension de "homme" car cet individu n'est pas en effet un représentant prototypique des hommes (Desclés, *Ibid.*).

Dans l'acception courante aujourd'hui, une algèbre booléenne est n'importe quel ensemble $\mathcal{B}$ muni des opérations binaires $*$ et $\circ$, de l'opération unaire $'$, dans lequel figurent les éléments distingués 1 et 0, et qui vérifie les axiomes suivants :

| | | |
|---|---|---|
| (1.54) | *Idempotence* | $x * x = x, \quad x \circ x = x$ |
| (1.55) | *Commutativité* | $x * y = y * x, \quad x \circ y = y \circ x$ |
| (1.56) | *Associativité* | $x * (y * z) = (x * y) * z$<br>$x \circ (y \circ z) = (x \circ y) \circ z$ |
| (1.57) | *Distributivité* | $x * (y \circ z) = (x * y) \circ (x * z)$<br>$x \circ (y * z) = (x \circ y) * (x \circ z)$ |
| (1.58) | *Existence d'un neutre* | $x * 1 = x, \quad x \circ 0 = x$ |
| (1.59) | *Existence d'un absorbant* | $x * 0 = 0, \quad x \circ 1 = 1$ |
| (1.60) | *Existence d'un complément* | $x * x' = 0, \quad x \circ x' = 1$ |

Pour obtenir la logique des classes à partir de l'algèbre de Boole, il suffit de spécifier ainsi l'ensemble et les opérations :

- L'ensemble $\mathcal{B}$ est l'ensemble des classes qui sont des extensions de termes généraux.
- L'opération $*$ est l'intersection ensembliste ($\cap$) : si $\alpha$ est la classe qui est l'extension du terme "ingénieur" et $\beta$ la classe qui est l'extension de "chef d'entreprise", $\alpha \cap \beta$ dénote la classe des êtres qui sont à la fois ingénieurs et chefs d'entreprise.
- L'opération $\circ$ est l'union ensembliste ($\cup$) : $\alpha \cup \beta$ dans l'exemple précédent est la classe dont les membres sont soit ingénieurs, soit chefs d'entreprise, soit les deux.
- 0 dénote la classe vide qui est l'extension de termes tels que "centaure", "licorne".
- 1 dénote la classe universelle, ce que De Morgan appelle l'univers du discours.
- L'opération $'$ sert à former la classe complémentaire ($-$). Si nous prenons comme univers l'ensemble de tous les individus existant soit physiquement soit mathématiquement, $-\alpha$ ou $\overline{\alpha}$, qui est une abréviation de $1-\alpha$, désigne la classe de tous les non-ingénieurs, classe qui contient notamment tous les nourrissons et tous les nombres pairs.
- Grâce à la notion d'intersection de classes et de classe vide, on peut définir

la relation $\subset$ :

$$\alpha \subset \beta \quad =_{def} \quad \alpha \cap \overline{\beta} = 0.$$

C'est la relation d'inclusion entre classes.

### 1.6.2. *Le carré logique révisé*

La transcription de la logique des termes interprétés en extension n'est pas neutre. Elle affecte la logique traditionnelle en rendant non valides certaines inférences immédiates et certaines formes de syllogisme qui étaient tenues pour valides par Aristote. Cet écart entre la logique traditionnelle des termes et sa transcription en algèbre booléenne ne doit pas être interprété comme la preuve que notre manière de raisonner diffère de celle d'Aristote et qu'il s'est trompé. Il s'agit en fait de tout autre chose. Si l'algébrisation de la logique des termes ne permet pas de justifier toutes les inférences reconnues valides par Aristote, cela tient au fait que la validité tient à l'interprétation que l'on donne aux constantes logiques. Or la traduction de la logique en algèbre de Boole passe par une altération sémantique des pronoms indéfinis "tous" et "aucun".

Dans l'usage ingénu de la langue naturelle, qu'il s'agisse du grec ancien utilisé par Aristote ou du français contemporain, l'énoncé "Tous les enfants de Philippe sont endormis" véhicule une présupposition, la *présupposition* que les enfants de Philippe existent. Avec Strawson (1952, p. 175) nous dirons qu'un énoncé $S$ présuppose un énoncé $S'$ si la vérité de $S'$ est une précondition de la vérité ou de la fausseté de $S$. "Il a cessé de fumer" présuppose "Il a fumé" dans le sens que Strawson donne au verbe *présupposer*. Si l'on prend "Tous les $A$ sont $B$" dans le sens qu'on vient de préciser, il est clair qu'on peut en inférer "Quelques $A$ sont $B$".

Au contraire, si on interprète "Tous les $A$ sont $B$" par l'équation "l'intersection de la classe des $A$ et de la classe des non-$B$ est vide" et "Quelques $A$ sont $B$" par l'équation "l'intersection de la classe des $A$ et de la classe des $B$ n'est pas vide (mais occupée par des individus)", il est clair qu'on ne pourra plus inférer la vérité de la particulière affirmative à partir de la vérité de l'universelle affirmative. En effet, dans la nouvelle acception de "Tous les $A$ sont $B$", si A est vide, ce qui sera le cas si $A$ est la classe des centaures, l'intersection de la classe $A$ et de la classe non-$B$ est vide aussi. La condition $\alpha \cap \overline{\beta} = 0$ est donc vérifiée et la proposition "Tous les centaures sont blancs", par exemple, est vraie. En revanche, la proposition "Quelques centaures sont blancs", $\alpha \cap \beta \neq 0$, n'est pas vraie, car si $\alpha$ est vide, il est faux que l'intersection de la classe $\alpha$ et de la classe $\beta$ soit distincte du vide.

L'*explication générale* de l'impossibilité d'inférer une particulière d'une universelle est la suivante : le fait que l'intersection de la classe $\alpha$ et de la classe non-$\beta$ est vide ne nous permet pas de savoir si la classe $\alpha$ est vide ou non. En revanche, le fait que l'intersection de la classe $\alpha$ et de la classe $\beta$ est distincte

du vide implique que la classe $\alpha$ n'est pas vide.

Avant de choisir entre l'interprétation aristotélicienne et l'interprétation booléenne de l'universelle, il faut s'accorder sur un critère : va-t-on privilégier l'usage courant ou consentira-t-on à modifier celui-ci à des fins scientifiques ? Il y a de bonnes raisons de préférer la deuxième attitude : adopter la lecture booléenne accroît le pouvoir expressif de la langue. En nous débarrassant des présuppositions d'existence qui paraissaient indissolublement liées à l'universalité, nous nous libérons de l'obligation d'exprimer deux idées à la fois (l'universalité et l'existence) et nous nous donnons la possibilité de les exprimer séparément, sans nous priver pour autant de la possibilité de les réunir à nouveau quand nous le voudrons. C'est un progrès car il y a des cas où nous voulons exprimer l'universalité *sans* affirmer ou présupposer l'existence d'individus auxquels s'applique le terme sujet. C'est le cas, par exemple, pour la phrase 1.61.

(1.61) Tous les corps sur lesquels ne s'exerce aucune force poursuivent leur course selon une trajectoire rectiligne.

La mécanique affirme la vérité de cet énoncé, sans préjuger de l'existence de corps sur lesquels ne s'exerce aucune force.

Pour représenter les relations entre classes existant dans l'algèbre booléenne, les diagrammes d'Euler ne suffisent plus. Il faut en effet représenter l'univers du discours et exprimer la différence entre trois possibilités : celle où on sait que la classe est vide, celle où on sait que la classe est non vide, et celle où on ne sait rien sur la classe. Des hachures représenteront le vide et on laissera en blanc la zone représentant la classe à propos de laquelle on n'a pas d'information. La présence d'une croix dans une zone indique que celle-ci est occupée (n'est pas vide).

Pour représenter les rapports entre deux classes, on utilisera les diagrammes de Venn (1881).

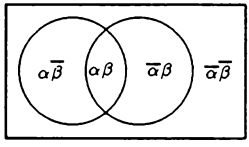



**Figure 1.13** *Classes et univers*

Les quatre propositions générales $A$, $E$, $I$, $O$ seront représentées respectivement par les quatre diagrammes des figures 1.14 à 1.17. Les hachures indiquent qu'une zone est vide. Les croix indiquent qu'une zone est occupée.

Dans le carré logique révisé (figure 1.18), seules les relations entre propositions qui sont la contradictoire l'une de l'autre subsistent (Eaton, 1931, p. 224).

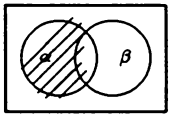



**Figure 1.14** *A (Affirmative universelle)*

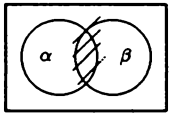



**Figure 1.15** *E (Négative universelle)*

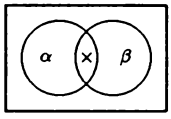



**Figure 1.16** *I (Affirmative particulière)*

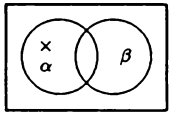



**Figure 1.17** *O (Négative particulière)*

### 1.6.3. *Le traitement algébrique des syllogismes*

Un syllogisme, comme tout raisonnement déductif, est *valide* si et seulement si sa forme est telle que toutes les interprétations qui rendent vraies ses

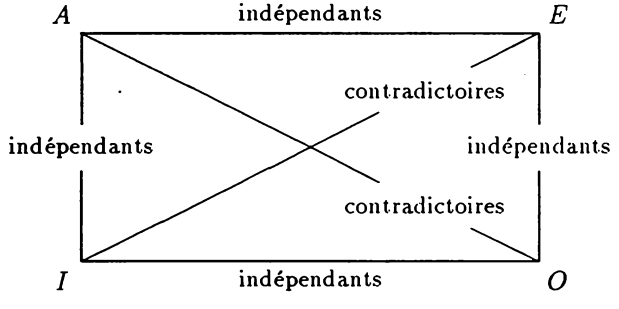



**Figure 1.18** *Carré logique selon Venn*

prémisses rendent également vraie la conclusion.

Aristote a formulé des règles permettant de déterminer si un syllogisme est valide ou non. Nous disposons aujourd'hui de méthodes algébriques et graphiques pour répondre à cette même question. La méthode algébrique est celle-ci : la conclusion d'un syllogisme découle validement des prémisses si et seulement si elle peut être obtenue par application des lois de l'algèbre de Boole.

Les prémisses d'un syllogisme sont deux universelles, deux particulières ou une universelle et une particulière. Dans le deuxième cas, le syllogisme n'est jamais valide. On peut donc se borner à étudier le premier (1) et le troisème (2) cas et à formuler les sept théorèmes de l'algèbre de Boole dont on a besoin pour déterminer si un syllogisme de l'espèce (1) ou de l'espèce (2) est valide ou non.

### 1.6.3.1. *Les deux prémisses sont des propositions universelles*

Dans ce cas, on a besoin des théorèmes 1.62 et 1.63 de l'algèbre de Boole :

(1.62) Si $\alpha = 0$ et $\beta = 0$ alors $\alpha \cup \beta = 0$,

(1.63) Si $\alpha x \cup \beta \overline{x} = 0$ alors $\alpha\beta = 0$.

*Exemple de syllogisme valide*

(1.64) Tous les hommes sont des animaux
Aucun animal n'est éternel
Donc aucun homme n'est éternel

Les prémisses sont $\alpha\overline{\beta} = 0$ et $\beta\gamma = 0$. Pour en tirer la conclusion $\alpha\gamma = 0$, on applique successivement le théorème 1.62 : $\alpha\overline{\beta} \cup \beta\gamma = 0$ et le théorème 1.63 ($\alpha\gamma = 0$). On obtient bien l'équation qui est la transcription algébrique de la conclusion. Le raisonnement est donc valide.

*Exemple de syllogisme non valide*

(1.65) Tous les $\alpha$ sont $\beta$
Tous les $\gamma$ sont $\beta$
Donc tous les $\alpha$ sont $\gamma$

Les prémisses sont $\alpha\overline{\beta} = 0$ et $\gamma\overline{\beta} = 0$. La conclusion est $\alpha\overline{\gamma} = 0$. Le théorème 1.62 donne $\alpha\overline{\beta} \cup \gamma\overline{\beta} = 0$. Il n'est pas possible d'appliquer le théorème 1.63, indispensable pour l'élimination du moyen terme $\beta$. Lorsque les deux prémisses sont particulières, le syllogisme n'est jamais valide. Reste donc à examiner le cas où l'une des prémisses est particulière.

1.6.3.2. *Cas où une des prémisses est universelle et l'autre particulière*

Dans le cas où l'une des prémisses est une universelle et l'autre une particulière, le raisonnement est parfois valide, parfois non valide. Ici aussi il est valide lorsque l'élimination du moyen terme peut être justifiée algébriquement. Dans le cas présent cependant, cette élimination exige l'exécution préalable de certaines transformations opérées sur les propositions initiales.

1. Il faut exprimer les données sous forme d'équations ou d'inéquations dont un membre est 0. On peut le faire à l'aide du théorème 1.66, dans lequel le symbole $\equiv$ se lit "si et seulement si".

(1.66) $$\alpha = \beta \quad \equiv \quad \alpha\overline{\beta} \cup \overline{\alpha}\beta = 0$$

2. Il faut ensuite faire en sorte que chacune des trois variables de classe figure dans la formulation des deux prémisses. On applique pour cela le théorème

(1.67) $$\alpha \quad = \quad \alpha(\beta \cup \overline{\beta}) \quad = \quad \alpha\beta \cup \alpha\overline{\beta}.$$

3. On a obtenu, grâce à (a), une équation et une inéquation. Si elles ont en commun un des termes de la somme, celui-ci peut être montré nul en vertu du théorème

(1.68) Si $\alpha \cup \beta = 0$, alors $\alpha = 0$ et $\beta = 0$.

4. Ce membre doit alors être éliminé au moyen du théorème

(1.69) Si $\alpha \cup \beta \neq 0$ et $\alpha = 0$ alors $\beta \neq 0$.

5. Enfin le moyen terme sera éliminé de la conclusion à l'aide du théorème

(1.70) Si $\alpha\beta \neq 0$, alors $\alpha \neq 0$.

Ces sept théorèmes peuvent être aisément démontrés à partir des axiomes de l'algèbre de Boole dus à Huntington.

*Exemple.*

(1.71) Tous les hommes sont des animaux
Quelques hommes sont très intelligents
donc quelques animaux sont très intelligents

Les classes utilisées sont : $\alpha$ = classe des hommes, $\beta$ = classe des animaux, $\gamma$ = classe des êtres très intelligents.

Les prémisses sont $\alpha\overline{\beta} = 0$ et $\alpha\gamma \neq 0$. La conclusion est $\beta\gamma \neq 0$. La formule $\alpha\overline{\beta} = 0$ devient, en vertu de 1.67, $\alpha\overline{\beta}(\gamma \cup \overline{\gamma}) = 0$, puis devient $\alpha\overline{\beta}\gamma \cup \alpha\overline{\beta}\overline{\gamma} = 0$ par distributivité; $\alpha\gamma \neq 0$ devient $\alpha\gamma(\beta \cup \overline{\beta}) \neq 0$, puis devient $\alpha\gamma\beta \cup \alpha\gamma\overline{\beta} \neq 0$ et, par commutativité, $\alpha\beta\gamma \cup \alpha\overline{\beta}\gamma \neq 0$.

On sait que $\alpha\overline{\beta}\gamma = 0$ en vertu de 1.68. On l'élimine de $\alpha\beta\gamma \cup \alpha\overline{\beta}\gamma \neq 0$ par 1.69 pour obtenir $\alpha\beta\gamma \neq 0$ et enfin, par élimination de $\alpha$ en vertu de 1.70, on arrive à $\beta\gamma \neq 0$. C.Q.F.D.

### 1.6.4. *La méthode des diagrammes de Venn*

#### 1.6.4.1. *Méthode*

Pour tester les syllogismes, il existe une méthode très intuitive : celle des diagrammes de Venn. Un syllogisme est valide lorsque ce que la conclusion affirme ne dépasse pas ce qui est affirmé par les prémisses. Pour s'en assurer, on note ce qui est affirmé par les prémisses au moyen de hachures pour les propositions universelles et de croix pour les propositions particulières. Si, pour exprimer ce qui est affirmé par la conclusion, on ne doit plus rien écrire, c'est que la conclusion n'affirme rien qui ne soit déjà affirmé par les prémisses et que le syllogisme est valide. Il est non valide dans le cas contraire.

*Remarques.*

1. Quand on utilise le diagramme pour tester un syllogisme qui possède une prémisse particulière et une universelle, il faut commencer par représenter l'universelle.

(1.72) Tous les aviateurs sont braves
Quelques aviateurs sont capitaines
donc quelques capitaines sont braves

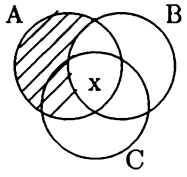



**Figure 1.19** *Syllogisme 1.72*

2. Il faut mettre la croix sur la ligne limitant deux secteurs lorsque le contenu affirmé dans les prémisses ne tranche pas la question.

(1.73) Tous les braves sont cléments
Quelques aviateurs sont cléments
donc quelques aviateurs sont braves

Si les prémisses sont vraies, la conclusion l'est peut-être, mais pas nécessairement. Pour transcrire la conclusion, il faut *déplacer* la croix vers l'intérieur du triangle situé à l'intersection des trois cercles, donc le raisonnement n'est pas valide : la conclusion affirme *plus* que les prémisses.

Le diagramme est une sorte de maquette qui permet de percevoir *de visu* les rapports entre classes affirmés par le syllogisme testé. On a en quelque sorte remplacé les classes concernant des objets éloignés par des classes de points situés dans une aire restreinte. Faut-il en conclure que l'épreuve que

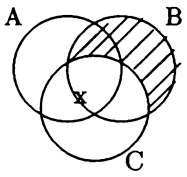



**Figure 1.20** *Syllogisme 1.73*

l'on fait subir à un syllogisme n'a de valeur que pour les classes particulières concernées ? Et qu'il faut une *induction*, une *généralisation*, pour conclure, partant de la validité du syllogisme concernant des surfaces observables, au syllogisme concernant les classes d'objets éloignés que ces surfaces symbolisent ? Non. Comme le fait observer Eaton, les *rapports* entre classes sont les mêmes, quelles que soient les classes. Or la validité dépend de ces rapports et non des classes elles-mêmes.

Le défaut de la méthode de Venn est qu'elle devient rapidement impraticable quand plus de trois classes entrent en jeu dans un raisonnement.

### 1.6.5. *Intension et extension des phrases*

La signification des termes généraux se scinde en intension et extension. Il en va de même pour la signification des phrases déclaratives. La théorie la plus courante assigne comme intension aux phrases déclaratives leurs *conditions de vérité* et comme extension leur *valeur de vérité*. Dans le *Tractatus logico-philosophicus*, Wittgenstein écrit : "une proposition est l'expression de ses conditions de vérité" (Wittgenstein, 1921, 4431). On peut connaître les conditions de vérité d'une phrase sans connaître sa valeur de vérité; par exemple, on peut connaître les conditions de vérité de la phrase "Le roi Arthur a existé" sans connaître ni la valeur de vérité de cette phrase, ni la manière de la vérifier ou même de la confirmer : le roi Arthur a existé si un être humain portant ce nom et détenteur de ce titre a vécu au lieu et à l'époque que les récits mentionnent. Il n'a pas existé s'il est simplement un personnage de fiction. Tout ceci a été clairement formulé dans le *Tractatus logico-philosophicus* :

> Comprendre une proposition signifie savoir ce qui est le cas si elle est vraie (on peut donc la comprendre, sans savoir si elle est vraie). Elle est comprise par quiconque comprend ses constituants (Wittgenstein, 1921, 4024).

Il y a des phrases qui ont plusieurs conditions de vérité différentes. C'est le cas des phrases ambiguës. Par exemple "les jumelles de l'opticien grossissent" est une phrase qui peut exprimer deux propositions différentes. Dans la première acception cette phrase est vraie *à condition que* des lunettes d'une

certaine espèce fassent voir les objets plus grands qu'on ne les voit à l'oeil nu. Dans la deuxième acception elle est vraie *à condition* que le marchand de lunettes ait des filles nées d'un même accouchement et qui prennent de l'embonpoint. La phrase présente à la fois une ambiguïté sémantique comme celle que l'on trouve dans "J'ai soigné un berger allemand" et une ambiguïté syntaxique comme celle qu'on trouve dans "couverture de brochure défraîchie" où "défraîchie" peut qualifier "brochure" ou "couverture". En effet, "jumelle" a deux sens lexicaux et "grossit" appartient à deux catégories syntaxiques (à celle des verbes transitifs et à celle des verbes intransitifs). Il faut éviter de telles phrases dans une inférence car elles conduisent facilement à des fautes de raisonnement connues sous le nom de *faute de l'équivoque*. Une telle faute est commise dans le raisonnement 1.74 qui, malgré les apparences, n'est pas un syllogisme (en effet, il contient *quatre* termes généraux et non pas trois : "grue" compte pour deux).

(1.74) Toutes les grues volent
Toutes les machines de levage et de manutention sont des grues
Donc toutes les machines de levage et de manutention volent

Il y a des phrases dont la *valeur de vérité* est instable. C'est le cas, par exemple, de la phrase "Il fait froid ce matin". Elles posent des problèmes au logicien qui a intérêt à les convertir en phrases intemporelles dans lesquelles la locution "ce matin" a été remplacée par une date.

### 1.6.6. *De l'algèbre de Boole à la logique des propositions*

Dans *The Mathematical Analysis of Logic*, Boole a montré comment on pouvait appliquer son algèbre des classes aux propositions. Il propose de donner comme extension à la phrase déclarative la classe des circonstances dans lesquelles elle est vraie. Si l'on exclut les phrases qui changent de valeur de vérité, il n'y aura que deux classes à considérer : la classe universelle et la classe vide.

On devra cependant ajouter le *principe de bivalence* affirmant que *toute proposition est vraie ou fausse*, ce qui, dans la lecture de Boole, devient : *Pour toute proposition, la classe des circonstances où elle est vraie est la classe universelle ou la classe vide*,

(1.75) Pour toute classe $\alpha$, $\alpha = 1$ ou $\alpha = 0$,

et donner comme membres à l'ensemble $\mathcal{B}$ non pas des classes quelconques, mais la classe des circonstances dans lesquelles les propositions $p, q, r$, etc... sont vraies. Tout le reste demeure inchangé.

A première vue, une difficulté se présente. Considérons les deux phrases 1.76 et 1.77.

(1.76) Tous les Grecs sont européens

(1.77) Ou $2+2=4$, ou $2+2\neq 4$

Traduite dans l'algèbre de Boole interprétée en logique des termes, la phrase 1.76 devient ($\alpha$ étant la classe des Grecs et $\overline{\beta}$ le complément de la classe des Européens) :

(1.78) $\alpha \cap \overline{\beta} = 0.$

Traduite dans l'algèbre de Boole interprétée en logique des propositions, la phrase 1.77 devient ($\gamma$ étant la classe des circonstances où "$2+2=4$" est vrai et $\overline{\gamma}$ le complément de cette classe) :

(1.79) $\gamma \cup \overline{\gamma} = 1.$

Il y a un lien très étroit entre les classes et les propositions : *toutes* les propositions qui sont des vérités de logique dans la logique des propositions deviennent en algèbre booléenne des classes dont on peut démontrer algébriquement qu'elles sont identiques à la classe universelle. Pour prouver qu'une proposition est une vérité de la logique des propositions, il suffira donc de la convertir en un foncteur de classe et de démontrer par l'algèbre de Boole que le nom en question désigne la classe universelle.

Prenons par exemple le principe suivant (cf. équation 1.29) de la logique des propositions (formulé à l'aide de *variables propositionnelles*, comme celles qui figurent dans les schémas d'inférence des Stoïciens, et non de *variables de classe*) :

(1.80) Que [si ($p$ et $q$) alors $r$]
implique que [si ($p$ et non $r$) alors non $q$].

Remplaçons par le signe $\supset$ les deux conditionnels, mais aussi le verbe "implique" qui, par abus de langage, tenait lieu d'un conditionnel :

(1.81) $[(p \wedge q) \supset r] \supset [(p \wedge \overline{r}) \supset \overline{q}].$

Eliminons les trois occurrences du conditionnel philonien $\supset$ au moyen de la définition

(1.82) $X \supset Y =_{def} \overline{X} \vee Y$ ("$\vee$" est le "ou" inclusif),

on obtient

(1.83) $\neg[\neg(p \wedge q) \vee r] \vee [\neg(p \wedge \overline{r}) \vee \overline{q}].$

Convertissons cet *énoncé* de la logique des propositions en un *foncteur* désignant une classe

(1.84) $-[-(\alpha \cap \beta) \cup \gamma] \cup [-(\alpha \cap \overline{\gamma}) \cup \overline{\beta}].$

Par les lois de De Morgan, 1.84 devient 1.85 :

(1.85) $[(\alpha \cap \beta) \cap \overline{\gamma}] \cup [(\overline{\alpha} \cup \gamma) \cup \overline{\beta}].$

Par associativité du $\cup$, on obtient

(1.86) $[(\alpha \cap \beta) \cap \overline{\gamma}] \cup (\overline{\alpha} \cup \gamma \cup \overline{\beta}).$

Par distributivité de $\cap$ sur $\cup$, on obtient

(1.87) $[(\alpha \cap \beta) \cup (\overline{\alpha} \cup \gamma \cup \overline{\beta})] \cap [\overline{\gamma} \cup (\overline{\alpha} \cup \gamma \cup \overline{\beta})]$.

Par associativité du $\cup$, on obtient

(1.88) $[(\alpha \cap \beta) \cup (\overline{\alpha} \cup \gamma \cup \overline{\beta})] \cap (\overline{\gamma} \cup \overline{\alpha} \cup \gamma \cup \overline{\beta})$.

Par distributivité du $\cap$ sur $\cup$, on obtient

(1.89) $[(\alpha \cup \overline{\alpha} \cup \gamma \cup \overline{\beta}) \cap (\beta \cup \overline{\alpha} \cup \gamma \cup \overline{\beta})] \cap (\overline{\gamma} \cup \overline{\alpha} \cup \gamma \cup \overline{\beta})$.

Sachant que $\alpha \cup \overline{\alpha} = 1$, et de même pour $\beta$ et $\gamma$,

(1.90) $(1 \cup \gamma \cup \overline{\beta}) \cap (1 \cup \overline{\alpha} \cup \gamma) \cap (1 \cup \overline{\alpha} \cup \overline{\beta})$,

d'où

(1.91) $1 \cap 1 \cap 1$,

et enfin

(1.92) 1.

On a donc bien

(1.93) $-[-(\alpha \cap \beta) \cup \gamma] \cup [-(\alpha \cap \overline{\gamma}) \cup \overline{\beta}] = 1$.

### 1.6.7. *Opération et relation*

Il importe cependant de ne pas confondre *l'algèbre des classes*, même réduite à *deux* classes (la classe universelle et la classe vide) avec l'algèbre des propositions où les variables prennent comme valeurs les *deux* valeurs Vrai et Faux. Dans la première interprétation de l'algèbre, la distinction entre *opération* ($-$, $\cap$, $\cup$) et *relation* ($\subset$, $=$) subsiste. Dans la deuxième interprétation de l'algèbre, c'est-à-dire dans l'algèbre propositionnelle ou algèbre des valeurs de vérité, la distinction disparaît :

> [...] maintenant chaque opération booléenne est en fait une relation (Scott, 1971, p. 792).

Pour comprendre la portée de cette pénétrante observation de D. Scott, comparons les six formules suivantes :

(1.94) $(1 + 2) < 4$

(1.95) $(1 < 2) < 4$

(1.96) $(\alpha \cup \beta) \subset \gamma$ c'est-à-dire $(\alpha \cup \beta) \cap \overline{\gamma} = 0$ (voir 1.6.1 loi 7)

(1.97) $(\alpha \subset \beta) \subset \gamma$ c'est-à-dire $\{[(\alpha \cap \overline{\beta}) = 0] \cap \overline{\gamma}\} = 0$ (idem)

(1.98) $(p \vee q) \supset r$

(1.99) $(p \supset q) \supset r$

On voit que les formules 1.94, 1.96, 1.98 et 1.99 sont bien formées, alors que les formules 1.95 et 1.97 sont mal formées. La formule 1.95 est mal formée parce que la relation $<$ ne peut relier que des *termes singuliers* et que $(1 < 2)$ n'est

pas un terme singulier, mais une proposition. La formule 1.97 est mal formée pour les mêmes raisons : l'opération $\cap$ ne *peut* prendre comme arguments que des classes, or $(\alpha \cap \overline{\beta}) = 0$ ne désigne pas une classe, mais une proposition. Au contraire, $(p \supset q) \supset r$ est bien formée car $\supset$ exprime *simultanément* une opération et une relation.

### 1.6.8. *Algèbre de Boole et logique*

Nous avons jusqu'ici indiqué les ressemblances et les différences entre l'interprétation du système abstrait de l'algèbre de Boole en calcul des classes et l'interprétation du même système en calcul des propositions. Il importe à présent de voir *techniquement* comment on peut interpréter le système abstrait de l'algèbre booléenne comme une algèbre des propositions.

On commencera par se donner un ensemble de propositions atomiques $\{p, q, r \ldots\}$ auxquelles on appliquera les connecteurs primitifs $\wedge$, $\vee$ et $\neg$.

On voudrait construire une algèbre booléenne des propositions à partir de ceux-ci en interprétant $*$ comme $\wedge$, $\circ$ comme $\vee$ et $'$ comme $\neg$ et en interprétant 1 comme signifiant "tautologie" (cf. section 1.4.6) et 0 comme signifiant "proposition contradictoire". Cependant, observe B. Partee,

> [...] cela ne peut se faire aussi directement car il n'existe pas une tautologie *unique* ($p \vee \neg p$, $q \vee \neg q$, $\neg(r \wedge \neg r)$ sont toutes des tautologies), ni une contradiction unique, et les autres opérations ne sont pas uniques non plus (le "complément" de $\neg p \wedge \neg q$ pourrait être $p \vee q$, $\neg(\neg p \wedge \neg q)$, ou n'importe quelle autre proposition logiquement équivalente). La première tâche à accomplir sera donc de grouper les propositions en *classes d'équivalence* sous la relation "est tautologiquement équivalent à" [...] Soit $A$ une proposition. On désignera par $\mathcal{E}(A)$ la classe de toutes les propositions tautologiquement équivalentes à $A$ (Partee, 1978, pp. 133–134).

Ceci posé, l'interprétation propositionnelle de l'algèbre des classes de Boole peut être donnée au moyen des six clauses suivantes (Partee, *ibid.*) dont on peut vérifier aisément qu'elles sont licites, c'est-à-dire qu'elles ne dépendent pas du choix de la proposition dans la classe considérée. Par exemple, la vérité de $A \equiv A'$ et de $B \equiv B'$ implique celle de $A \vee B \equiv A' \vee B'$.

- Soit $\mathcal{B}$ l'ensemble des classes d'équivalence de propositions formées à l'aide des propositions atomiques d'un ensemble quelconque $\mathcal{P}$ de propositions atomiques.
- Définissons l'opération $\cup$ sur $\mathcal{B}$ par

$$\mathcal{E}(A) \cup \mathcal{E}(B) \quad = \quad \mathcal{E}(A \vee B).$$

- Définissons l'opération $\cap$ sur $\mathcal{B}$ par

$$\mathcal{E}(A) \cap \mathcal{E}(B) \quad = \quad \mathcal{E}(A \wedge B).$$

- Définissons $-$ par

$$-(\mathcal{E}(A)) \quad = \quad \mathcal{E}(\neg A).$$

- 1 est la classe d'équivalence de toutes les tautologies, c'est-à-dire

$$1 \quad = \quad \mathcal{E}(p \vee \neg p).$$

- 0 est la classe d'équivalence de toutes les contradictions, c'est-à-dire

$$0 \quad = \quad \mathcal{E}(p \wedge \neg p).$$

Les mêmes lois de l'algèbre booléenne nous permettent, à condition d'adopter l'interprétation qui convient, de démontrer les vérités de la logique des termes tels que

(1.100) $[(\beta \cap \overline{\gamma} = 0) \wedge (\alpha \cap \beta \neq 0)] \supset [\alpha \cap \gamma \neq 0]$,

que Boole appelle *primary propositions*, et les vérités de la logique des propositions telles que

(1.101) $[(p \wedge q) \supset r] \supset [(p \wedge \overline{r}) \supset \overline{q}]$,

que Boole appelle *secondary propositions*.

### 1.6.9. *Evaluation de l'apport de Boole*

Si Boole a réuni la logique aristotélicienne et la logique stoïcienne sous les lois d'une seule et même algèbre, c'est au prix d'une division au niveau de l'interprétation qui ne permet pas de dégager le lien qui les unit. Comme l'écrit Frege,

> Chez Boole les deux parties [la logique des concepts et la logique des jugements] se font suite en sorte que l'une est comme l'image en miroir (*Spiegelbild*) de l'autre, et pour cette raison même, il n'existe entre elles aucun lien organique (Frege, *Nachgelassene Schriften*, cité par Cl. Imbert, 1971a, p. 25 en note).

En effet, chez Boole, on doit obligatoirement *choisir* entre les deux interprétations décrites à la figure 1.21.

| $\mathcal{B}$ | Ensemble de classes $(\alpha, \beta, \gamma, \ldots)$ | Ensemble de propositions $(p, q, r, \ldots)$ |
|---|---|---|
| 1 | classe universelle (1) | Tautologie (**V**) |
| 0 | classe vide (0) | Contradiction (**F**) |
| $+$ | union de classes $(\cup)$ | disjonction de propositions $(\vee)$ |
| $\times$ | intersection de classes $(\cap)$ | conjonction de propositions $(\wedge)$ |
| $-$ | complément de classe $(-)$ | négation de proposition $(\neg)$ |
| $\subset$ | inclusion de classe $(\subset)$ | implication matérielle $(\supset)$ |
| $=$ | identité entre classes $(=)$ | équivalence matérielle $(\equiv)$ |

**Figure 1.21** *Interprétations de l'algèbre de Boole*

La logique de Boole a deux faiblesses que Frege a relevées et que Cl. Imbert formule en ces termes :

> la première est que les principes de la logique aristotélicienne n'ont pas été altérés : la logique de Boole a les mêmes limites que celle du Stagirite[5]. Elle n'a pas d'expression pour les jugements d'existence. Et elle n'a pas non plus les moyens de construire de nouvelles expressions conceptuelles et doit supposer que les concepts sont donnés (Imbert, 1971a, p. 23).

Quant à la seconde faiblesse,

> [elle] tient à ce que les symboles servant à représenter en une équation les rapports entre termes (ce que Boole appelle une *primary proposition*) servent également à représenter le rapport entre des énoncés propositionnels (*secundary proposition*). Pour ce faire, Boole identifie un énoncé à la classe des instants pendant lesquels cet énoncé est vrai. Outre qu'un symbolisme recevant, dans le cours d'un même calcul, deux interprétations est ambigu, Frege objecte qu'on ne peut écrire *simultanément* le contenu des propositions et le lien entre propositions comme il se fait dans la continuité du raisonnement mathématique. L'objection signifie que si l'on peut, à la rigueur, transcrire dans le symbolisme booléen la langue parlée du mathématicien, il faudra le faire en deux étapes. On écrira successivement les équations correspondant au rapport des termes, puis celles qui traduisent le rapport des propositions (Imbert, *Ibid*, p. 24).

L'*ambiguïté* que dénonce Cl. Imbert est bien réelle. Dans l'interprétation en logique des termes, l'inclusion de classe est traitée comme une *relation* ("$\alpha \subset \beta$" est une abréviation de "$\alpha \cap \overline{\beta} = 0$") tandis que dans l'interprétation en logique des propositions, elle est une *opération* ("$p \supset q$" est une abréviation de "$\overline{p \wedge \overline{q}}$").

Supposons que nous voulions transcrire en algèbre de Boole l'énoncé

(1.102) S'il n'y a pas de licornes, toutes les licornes sont blondes,

nous devons faire cette transcription en deux étapes. Premièrement,

(1.103) Si $\lambda = 0$ alors $(\lambda \cap \overline{\beta}) = 0$,

où $\lambda$ est la classe des licornes et $\overline{\beta}$ le complément de la classe des objets blonds. Nous devrons *ensuite* représenter le rapport entre les propositions $\lambda = 0$ et $\lambda \cap \overline{\beta} = 0$, que nous désignerons respectivement par $p$ et $q$. En d'autres termes, nous devrons mettre en équation

(1.104) $p \supset q$.

Soit $\gamma$ la classe des circonstances où $p$ est vraie et $\overline{\delta}$ le complément de la classe des circonstances où $q$ est vraie. La formule 1.104 se transcrit à l'aide de l'équation

(1.105) $(\gamma \cap \overline{\delta}) = 0$.

[5] Aristote vivait à Stagire en Macédoine au quatrième siècle avant notre ère.

On voit qu'on a eu besoin de *deux* étapes *successives*. On ne pourrait donc pas représenter un raisonnement dont la validité repose à la fois sur les rapports entre propositions et entre termes, tels que le raisonnement 2.36 de la section 2.1.7.

# Chapitre 2
# Les notions fondamentales de la logique : de Frege à Gödel

## 2.1. La logique de Frege

### 2.1.1. *Utilité d'une langue artificielle*

Pour que nous puissions juger de la correction des preuves qui nous sont offertes dans une science déductive, il faut que nous connaissions toutes les propositions sur lesquelles l'auteur de la déduction s'appuie et toutes les règles qu'il se donne.

Or dans les preuves conduites dans la langue naturelle, il est fréquent que l'on fasse des raisonnements incomplets — ce que l'on appelle *enthymèmes* — dans lesquels certaines prémisses sont sous-entendues. Frege dénonce en ces termes ce défaut de la langue naturelle :

> Les formes où s'exprime la déduction sont si diverses, si lâches, si mal définies, que des hypothèses peuvent être introduites sans qu'on y prenne garde, et on omet de les compter quand on récapitule les conditions nécessaires à la validité de la conclusion [...] Même un auteur aussi scrupuleux et rigoureux qu'Euclide fait usage à plusieurs reprises d'hypothèses tacites [...] (Frege, 1882, 1971a, p. 65).

En outre, les règles d'inférence internes à la langue naturelle (comme par exemple "De $p$ et $q$ on peut déduire $p$ et on peut déduire $q$, en raison du sens de "et" ") ne forment pas un ensemble aux contours nets :

> Le langage n'offre pas un lot bien délimité de formes de déduction et, à s'en tenir à la forme linguistique, il n'est pas possible de distinguer une séquence sans lacune de celle qui omet des propositions intermédiaires (Frege, *Ibid.*, p. 65).

Enfin la langue naturelle nous induit en erreur. Parlant du langage naturel, Frege écrit :

> Il ne satisfait pas à la condition ici primordiale, celle d'univocité. Les cas les plus dangereux sont ceux où les significations des mots diffèrent très peu, où les variations sont légères bien que non équivalentes. Parmi de nombreux exemples on citera un cas typique fort commun : c'est le même mot qui sert à désigner un concept et un objet particulier tombant sous ce concept; de manière générale, aucune différence n'est marquée entre le concept et l'objet particulier. "Le cheval" peut désigner un individu mais tout aussi bien l'espèce, comme dans la proposition "le cheval est un herbivore"; et "cheval" peut enfin avoir le sens d'un concept comme dans la proposition "ceci est un cheval". La langue n'est pas régie par des lois logiques telles que l'observance de la grammaire puisse suffire à garantir la rigueur formelle du cours de la pensée (Frege, *Ibid.*, p. 64).

On pourrait illustrer le propos de Frege par le raisonnement non valide 2.1, raisonnement dont la non-validité est masquée par la langue naturelle :

(2.1) Les Indiens d'Amérique sont en voie d'extinction
Oeil de faucon est un Indien d'Amérique
donc Oeil de faucon est en voie d'extinction.

L'erreur vient de ce que "Indiens d'Amérique" s'applique à des tribus dans sa première occurrence et à des individus dans sa seconde occurrence. Pour remédier à cette situation, Frege a entrepris la construction d'une langue artificielle où seraient codifiées *toutes* les règles d'inférence et les lois logiques dont on a besoin pour la conduite des preuves en mathématique. Il ne s'agit pas d'introduire localement de la rigueur comme l'avaient fait Boole et Schröder, mais de supplanter la langue naturelle pour ce qui est de la conduite du raisonnement[1].

Dans cette langue artificielle, Frege formulera les lois logiques. Comme celles-ci sont en nombre infini, il n'est pas envisageable de les énumérer. D'ailleurs, une telle énumération, même si elle était possible, ne nous révélerait pas les connexions qui existent entre les lois. Pour dominer l'ensemble infini des lois logiques nécessaires aux preuves mathématiques, Frege adopte le mode déductif de présentation :

> De cette manière nous arrivons à un petit nombre de lois [en fait, il n'aura besoin que de neuf axiomes] dans lesquelles, si nous ajoutons celles qui sont contenues dans les règles, le contenu de toutes les lois est inclus, quoique dans un état non déployé. Et le fait que le mode déductif de présentation nous fasse connaître ce noyau est un autre de ses

[1] On pourrait objecter que la langue artificielle ne fait que déplacer le problème. Les difficultés surgiront au moment de l'encodage en langue artificielle des raisonnements concrets de la langue naturelle, mais il ne faut pas surestimer cette objection.

> avantages. Comme l'infinie multitude des lois énonçables nous empêche de les énumérer toutes, nous ne pouvons atteindre la complétude qu'en isolant celles, parmi ces lois, qui *par leur pouvoir* les contiennent toutes (Frege, 1879, 1967, p. 29).

La langue artificielle que Frege a construite pour la conduite des preuves est un *calcul* au sens de Leibniz dans lequel, comme le note Cl. Imbert,

> [...] l'écriture des formules et l'inférence d'une formule à partir d'une autre sont soumises à des règles peu nombreuses et ne souffrant aucune exception, en sorte que la conclusion d'un raisonnement est établie par la seule considération des symboles (Imbert, 1971a, p. 20).

Mais cette langue artificielle n'est pas seulement un calcul. C'est aussi ce que l'on appelle aujourd'hui une *notation canonique*, c'est-à-dire une *langue* dans laquelle les symboles ont un sens, même si celui-ci n'est pas utilisé dans la conduite des raisonnements. L'idéographie (*Begriffsschrift*) de Frege doit pouvoir *exprimer* aisément les notions mathématiques que la langue naturelle ne peut exprimer qu'avec peine (comme la notion de continuité d'une fonction).

### 2.1.2. *Les propriétés d'une langue idéale*

La langue naturelle ne sert pas seulement à exprimer le contenu d'une science impersonnelle et à conduire des preuves, elle sert aussi à communiquer et à dialoguer avec un interlocuteur. Elle comporte des mots qui n'ont rien à voir avec les *conditions de vérité*, ni avec la *valeur de vérité* des phrases, mais avec la dimension communicative du langage :

> Pensons, écrit Frege, à des mots comme "encore", "déjà". Dans la proposition "Alfred n'est pas encore venu", on dit "Alfred n'est pas venu" et on y indique que l'on attend sa venue, mais on l'indique seulement. On ne peut pas dire que le sens de la proposition soit faux s'il se trouve que la venue d'Alfred n'est pas attendue (Frege, 1918–1919, 1971a, p. 177).

Frege donne également l'exemple de "mais" qui a été récemment étudié par O. Ducrot (voir Jayez, 1988). Comparons les propositions 2.2 et 2.3 :

(2.2) Il fait mauvais *et* nous irons nous promener

(2.3) Il fait mauvais, *mais* nous irons nous promener

Les deux phrases ont les mêmes *conditions de vérité*. Ce que la deuxième exprime *en plus* par l'emploi de la conjonction "mais", c'est que les deux phrases liées par ce mot ont une direction argumentative opposée. "Il fait mauvais" constitue, en effet, un argument possible pour "nous n'irons pas nous promener".

La langue idéale pour la formulation de la science et la conduite des preuves devra éliminer ces traits de la langue naturelle, ces surdéterminations du *sens* des phrases (conçu comme *conditions de vérité*). Il faudra également bannir les

pronoms personnels comme "je", "tu", les circonstanciels égocentriques ("ici", "maintenant"), les temps du verbe etc.

Plus fondamentalement, Frege va libérer la logique de la tutelle de la grammaire en remplaçant l'analyse de la proposition en sujet et prédicat par l'analyse en fonction et arguments. Il s'agit là d'une rupture majeure avec la tradition aristotélicienne qui avait survécu même chez Boole. Considérez, nous dit Frege, les deux propositions "les Grecs ont défait les Perses à Platée" et "Les Perses ont été défaits par les Grecs à Platée". Ces deux propositions ont même contenu conceptuel. Elles ne diffèrent que par la manière dont sont remplies les places du sujet et du prédicat. Or la distinction entre sujet et prédicat n'a de rôle que dans la dimension interlocutoire du langage :

> S'agissant du langage ordinaire, la place traditionnellement réservée au sujet dans la séquence verbale n'est rien d'autre qu'une place *distinguée*, nous y mettons ce vers quoi nous souhaitons spécialement diriger l'attention de l'auditeur (Frege, 1879, 1967, p. 12, cité par F. Jacques, 1979, p. 708).

Frege entend ne pas reproduire dans son idéographie, sa langue artificielle, la distinction sujet-prédicat; la raison, c'est qu'il entend ne prendre en considération que les différences dans la proposition "qui influencent ses *conséquences possibles*", c'est-à-dire son potentiel d'inférence.

Remarquons ici que si Frege avait considéré non seulement les inférences dont on fait usage en mathématiques, mais aussi celles qui portent sur des états psychologiques, la distinction entre voix active et voix passive serait devenue pertinente pour la logique. L'exemple suivant, dû à S. Ben Chorin le montre bien. Quoique les deux propositions atomiques suivantes qui ne diffèrent que par la *voix*, "Marie embrasse le bébé" et "le bébé est embrassé par Marie", aient le même contenu conceptuel à l'état isolé, il se fait qu'elles acquièrent un sens différent quand elles figurent dans une proposition composée, à titre de complément d'un verbe (ou d'une expression) désignant un acte mental dirigé sur un objet intentionnel ("croire", "dériver", "craindre") :

(2.4) Jean est jaloux de ce que (...)

L'expression "est jaloux de ce que" agit comme un prisme sur un rayon lumineux. C'est le même *état de chose* qui est décrit par la phrase à la voix active et par la phrase à la voix passive, mais l'opposition *topic-comment* n'est pas la même : dans un cas, l'individu dont on parle est Marie, dans l'autre, c'est le bébé. Cette différence est imperceptible tant que la phrase est utilisée isolément. Elle devient perceptible dès que celle-ci est enchâssée. Le "budget inférentiel" de la phrase active se révèle différent de celui de la phrase passive. Du fait que Jean est jaloux de ce que Marie — plutôt que lui — embrasse le bébé, il ne s'ensuit pas que Jean soit jaloux de ce que le bébé — plutôt que lui — soit embrassé par Marie.

### 2.1.3. *Assertion et contenu propositionnel*

Dans la logique d'Aristote, les prémisses et les conclusions des raisonnements sont appelées *propositions*. Frege introduit une distinction importante entre le contenu propositionnel et l'acte d'assertion.

Le contenu propositionnel non asserté avait déjà été isolé dans l'antiquité sous le nom de *lexis*. La lexis est définie en ces termes : "énoncé susceptible d'être dit vrai ou faux, mais qui n'est considéré que dans son contenu, et sans affirmation ni négation actuelle" (Lalande, 1951, p. 557). Goblot introduit la locution *jugement virtuel* pour désigner le contenu propositionnel :

> les jugements virtuels [...] sont des jugements complets : ils ont leur sujet, leur attribut, leur copule, tous leurs caractères formels; il ne leur manque que la croyance (*Ibid.*).

L'acte d'assertion est défini ainsi dans le dictionnaire de Lalande :

> acte de l'esprit qui déclare vraie une proposition, que celle-ci soit, dans sa forme, affirmative [...] ou négative (83).

"Asserter" n'est donc pas un synonyme superflu de "affirmer". En effet, affirmer s'oppose à nier, tandis qu'asserter s'oppose à considérer un contenu propositionnel sans l'affirmer ni le nier. Il faut distinguer l'acte de *dénégation* (*denial* en anglais) de l'acte d'assertion portant sur un contenu propositionnel négatif. La négation, par opposition à la dénégation, fait partie du contenu propositionnel. On peut contempler un contenu propositionnel négatif, sans l'affirmer ni le nier.

La logique traditionnelle classait les jugements en jugements affirmatifs et négatifs, particuliers et universels, catégoriques, hypothétiques et disjonctifs. Toutes ces distinctions sont déplacées par Frege. Elles ne s'appliquent plus aux jugements, mais à leur contenu propositionnel.

Frege n'a pas inventé la distinction entre acte d'assertion et lexis, mais il en a révélé un aspect nouveau qui a un rôle décisif en logique : c'est le contenu propositionnel qui doit être investi du rôle de porteur d'une *valeur de vérité*, tandis que c'est le jugement ou proposition assertée qui joue le rôle de *prémisse* et de *conclusion* dans un raisonnement.

L'analyse de Frege est indispensable. Elle permet de comprendre le fonctionnement d'un raisonnement comme le *modus ponens*. Dans le *modus ponens*, la première prémisse est de forme hypothétique : ce qui est asserté, c'est une proposition conditionnelle de la forme "Si $p$ alors $q$", mais ni l'antécédent représenté ici par le marque-place "$p$", ni le conséquent "$q$" ne sont assertés. En revanche "$p$" et "$q$", bien que non assertés, ont une valeur de vérité. On ne comprend pas le *modus ponens* si on ne fait pas la distinction de Frege. En effet, si on dit que même l'antécédent et le conséquent sont assertés dans "Si $p$ alors $q$", le raisonnement ne sert plus à rien, il n'y a pas *progrès* de la pensée. Si d'autre part, on assigne une différence de contenu à une proposition assertée

et à une proposition non assertée, le *modus ponens* devient un raisonnement non-valide, les marque-place récurrents sont employés de manière *équivoque*.

Nous avons dit que les prémisses et la conclusion d'un raisonnement étaient des jugements, des propositions assertées. Cela vaut pour les *preuves catégoriques* qui peuvent, d'ailleurs, contenir des propositions conditionnelles. Nous verrons plus loin qu'il existe une forme de preuve, appelée preuve hypothétique, dans laquelle on part de prémisses qui ne sont pas vraiment assertées, mais plutôt posées ou supposées provisoirement. Ceci ne remet pas en cause la distinction entre asserté et non asserté, mais apporte une restriction à la thèse selon laquelle les prémisses sont toujours des propositions assertées.

### 2.1.4. *Fonction propositionnelle et proposition singulière*

Bien que l'*idée* de fonction propositionnelle remonte à Frege, la *locution* "fonction propositionnelle" a été introduite par Russell et Whitehead dans les *Principia mathematica*. Leur définition dans l'Introduction de la deuxième édition (1925) est la suivante :

> Soit $\phi x$ un énoncé contenant une variable $x$ et tel qu'il devienne une proposition quand $x$ reçoit une signification déterminée. Donc $\phi x$ est appelée une "fonction propositionnelle"; ce n'est pas une proposition puisque, dû à l'ambiguïté de $x$, il ne fait pas une assertion. Ainsi "$x$ est blessé" n'asserte rien tant que nous n'avons pas fixé qui est $x$ (Russell et Whitehead, 1925, p. 14).

La fonction propositionnelle est donc, en fait, une expression linguistique.

Frege distingue les propositions qui attribuent des propriétés à des individus de celles qui expriment des relations entre des individus. Les premières sont des fonctions propositionnelles à une place combinées à un terme singulier, les secondes, les fonctions propositionnelles à plusieurs places. La fonction propositionnelle à plusieurs variables permet d'analyser de manière satisfaisante les énoncés de relation sur lesquels achoppait l'analyse traditionnelle en sujet et prédicat. Leibniz, qui voulait s'en tenir à l'analyse du jugement singulier en sujet et prédicat, comprenant qu'on ne peut analyser

(2.5) Pâris aime Hélène

en sujet "Pâris" et prédicat "aime Hélène" car cela revient à conférer à l'un des termes singuliers un privilège que rien ne justifie, recourt à la prédication double et analyse ainsi la phrase 2.5 :

(2.6) Pâris est amant en tant qu'Hélène est aimée.

La paraphrase 2.6 est manifestement inadéquate, elle pèche par le manque d'analyse sémantique de l'opérateur *quatenus* c'est-à-dire *en tant que* (Blanché, 1970, p. 200). Au contraire, l'analyse de ces constructions grammaticales à l'aide de la construction *fonction + argument* permet de dégager aussi bien

les ressemblances que les différences entre les propositions attributives et les propositions relationnelles. La proposition

(2.7) Pâris chante

sera formée de la fonction propositionnelle "$x$ chante" et du terme singulier "Pâris", substitué à "$x$". La proposition

(2.8) Pâris aime Hélène

sera formée de la fonction propositionnelle "$x$ aime $y$" et des termes singuliers "Pâris" et "Hélène" substitués respectivement à $x$ et à $y$. Remarquons que le rôle prépondérant du sujet disparaît si on note "$x$ aime $y$" par "aime($x$,$y$)".

Certains grammairiens proposent une troisième sorte d'analyse, selon laquelle le verbe transitif "aime" suivi du nom "Hélène" devient un syntagme verbal "aime Hélène" appartenant à la catégorie syntaxique des verbes intransitifs et le syntagme combiné au nom "Pâris" (ou à un syntagme nominal) donne naissance à une phrase "Pâris aime Hélène".

L'analyse est peut-être pleinement satisfaisante pour le grammairien soucieux de rendre compte de la grammaticalité des phrases, c'est-à-dire de leur statut de phrase bien formée, mais pour le logicien qui cherche à dégager la forme logique des énoncés, l'analyse en termes de concaténation de mots ou de syntagmes est moins riche d'information que l'analyse frégéenne en termes de fonction propositionnelle et de variables sur lesquelles on peut opérer des substitutions. Pour le voir, considérons les trois phrases 2.9–2.11. L'exemple est emprunté à M. Dummett :

(2.9) Brutus tua Brutus,

(2.10) Cassius tua Cassius,

(2.11) Brutus tua Cassius.

Si on ne tient compte que des mots, les phrases 2.9 et 2.10 n'ont rien d'autre en commun que le fait de puiser l'une et l'autre leur lexique dans le lexique de la troisième. Si, au contraire, on se représente les propositions singulières comme construites par substitution opérée sur des fonctions propositionnelles, on constate que par surcroît, les phrases 2.9 et 2.10 se ressemblent entre elles et diffèrent de 2.11 en cela qu'elles ont en commun la fonction propositionnelle "$x$ tua $x$".

### 2.1.5. *Fonction propositionnelle et proposition générale*

La notion de fonction propositionnelle permet d'unifier des parties différentes de la logique tout en respectant leur intégrité. Elle a permis d'unifier la logique des propositions singulières attributives et celle des propositions singulières relationnelles sans pour cela réduire les secondes aux premières. Elle permet aussi d'unifier la logique des propositions singulières et celle des propo-

sitions générales, tout en reconnaissant en même temps la différence profonde qui les sépare, différence méconnue par la logique traditionnelle.

Considérons les deux propositions suivantes :

(2.12) Le nombre 20 est représentable par la somme de quatre carrés

(2.13) Chaque entier positif est représentable par la somme de quatre carrés

Selon l'analyse traditionnelle, elles ont le même prédicat, à savoir "est représentable par la somme de quatre carrés", elles ne diffèrent que par le sujet. Dans l'analyse de Frege, au contraire, la première proposition est une assertion au sujet du nombre 20, tandis que la deuxième proposition est une assertion au sujet de la fonction propositionnelle "$x$ est représentable par la somme de quatre carrés".

La *forme logique* de la proposition 2.12 peut être décrite ainsi : la proposition 2.12 s'obtient en substituant le terme singulier "le nombre 20" à "$x$" dans la fonction propositionnelle "$x$ est représentable par la somme de quatre carrés". La *forme logique* de la proposition 2.13 peut être décrite ainsi : la proposition 2.13 s'obtient en substituant la fonction propositionnelle "$x$ est représentable par la somme de quatre carrés" à "$X$" dans la fonction propositionnelle "$X$ est possédée par un entier positif arbitraire".

Comme le note Cl. Imbert, pour Frege la proposition

(2.14) Dieu existe

"ne dit pas d'un objet, Dieu, qu'il a la propriété d'exister. Elle dit d'un concept, *être Dieu*, qu'il n'est point vide". Frege écrivait que l'existence est une propriété de second niveau (Imbert, 1971a, p. 40).

Ces analyses attribuent à des phrases dont la *forme grammaticale* est apparemment assez proche une *forme logique très différente*. La justification de cette analyse se trouve dans son pouvoir explicatif : si on admet l'analyse de Frege, on est en mesure de rendre compte de la validité ou de la non-validité de tous les raisonnements construits avec les expressions "tous", "quelques", alors que la logique classique d'Aristote ne permettait de rendre compte que de la validité d'un petit nombre de ceux-ci. Par exemple, ni la logique d'Aristote, ni l'algèbre de Boole ne peut rendre compte de la validité des raisonnements 2.15 et 2.16.

(2.15) Il y au moins un club dont tous sont membres, donc tous sont membres d'au moins un club

(2.16) La relation $x > y$ est transitive, $4 > 3$ et $3 > 2$, donc $4 > 2$

Dans l'analyse frégéenne des propositions générales, la fonction propositionnelle occupe la place de ce qu'on appelait "sujet" dans l'analyse *sujet-prédicat*. Cette analyse a été reprise par Russell dans son célèbre article "On Denoting", dans lequel il expose pour la première fois sa théorie des descriptions définies (termes singuliers de la forme "l'individu $x$ tel que $x$ est $\phi$").

Russell propose en effet de représenter la proposition

(2.17) Tout est pesant

par

(2.18) "$x$ est pesant" est toujours vrai.

Il propose, en effet, l'interprétation suivante :

- $C$(tout) signifie "$C(x)$ est toujours vrai";
- $C$(rien) signifie " "$C(x)$ est faux" est toujours vrai";
- $C$(quelque chose) signifie "Il est faux que "$C(x)$ est faux" est toujours vrai"

(Russell, 1905, 1956, p. 42; tr. fr. *Philosophie* no 8, 1985, pp. 11–28).

Russell interprète "Tous les hommes sont mortels" comme signifiant " "Si $x$ est humain, alors $x$ est mortel" est toujours vrai" et "quelques hommes sont vivants" comme signifiant " "$x$ est un homme et $x$ est vivant" n'est pas toujours faux".

On serait tenté d'objecter que ces interprétations ne sont pas des interprétations éclairantes. En effet, elles n'*éliminent* pas les termes indéfinis "tous", "quelques", elles ne font que les *déplacer*. "Tous" qui figurait dans "tous les hommes sont mortels" se retrouve dans "toujours". Une telle critique manque de pertinence car l'intérêt de ces analyses est ailleurs. Il est dans la mise au jour d'une forme logique différente de la forme grammaticale, mise au jour qui se justifie, comme nous l'avons dit plus haut, par le fait qu'il sera possible, grâce à elle, d'exprimer toute la force logique (le potentiel d'inférence) de ces expressions.

La notation qu'on vient de présenter, notation dans laquelle les indéfinis "tous", "quelque", "aucun" deviennent des sortes de prédicats appliqués aux fonctions propositionnelles, est peu maniable. Si nous cherchons à l'appliquer aux deux propositions 2.19 et 2.20 — l'exemple est dû à Quine (1941, 1966, p. 17) — nous aboutissons à une paraphrase incompréhensible.

(2.19) Tous ceux qui intéressent tous les hommes intéressent Socrate

(2.20) Tous ceux que tous les hommes intéressent intéressent Socrate

Si l'on adopte la paraphrase de Russell, on obtient respectivement

(2.21) "Si ("Si $y$ est homme, $x$ intéresse $y$" est toujours vrai) alors $x$ intéresse Socrate" est toujours vrai,

(2.22) "Si ("Si $y$ est homme, $y$ intéresse $x$" est toujours vrai) alors $x$ intéresse Socrate" est toujours vrai.

Cette paraphrase est du charabia. Ce qui nous manque, c'est un *marqueur de portée* pour la généralité. C'est ici que Frege et Peirce, indépendamment l'un de l'autre, font une invention décisive dont Whitehead et Russell tireront pleinement parti dans les *Principia mathematica*, l'invention du quantificateur.

### 2.1.6. *L'introduction du quantificateur*

Bien que les articles où Peirce publie sa découverte soient légèrement postérieurs à la *Begriffsschrift*, c'est l'apport de Peirce que nous évoquerons en premier lieu parce que Peirce se situe dans le prolongement de Boole, plutôt qu'en rupture avec lui, comme Frege.

Peirce cherchait à étendre l'algèbre logique de Boole pour couvrir aussi les relations. Comme le rappelle G. H. Moore,

> C'est en appliquant les opérations de réunion et d'intersection aux relations que Peirce [1883] developpa la notion d'un quantificateur comme quelque chose de distinct des connecteurs booléens. (Moore, 1988, p.103)

Peirce observe qu'il est possible d'exprimer l'énoncé

(2.23) Tous les A sont B,

dans le langage de l'algèbre des classes, à l'aide de

(2.24) $\alpha \subset \beta$,

mais non l'énoncé

(2.25) Tous aiment quelqu'un.

Pour exprimer des énoncés de ce type, Peirce introduit deux nouveaux opérateurs ("$\Pi$" pour "tous" et "$\Sigma$" pour "quelque"). L'énoncé 2.25 sera représenté par

(2.26) $\Pi_i \Sigma_j \text{aime}(i, j)$

Dans un article légèrement postérieur, Peirce rapproche le quantificateur "Tous" du produit et le quantificateur "quelque" de la somme.

> [...] nous pourrions utiliser $\Sigma$ pour *quelque*, suggérant une somme, et $\Pi$ pour *tous*, suggérant un produit. [...] Il convient d'observer que $\Sigma_i x_i$ et $\Pi_i x_i$ sont seulement *semblables* à une somme et à un produit. A strictement parler, ils ne sont pas de cette nature parce que les individus de l'univers peuvent être innombrables. (Peirce, 1885, p.194-195, cité par Moore, 1980, p.104)

Frege, au contraire, introduit le quantificateur universel comme une notion primitive. A partir du quantificateur universel $\forall x$ (ou $\forall y$ etc.) qui se lit aussi "Pour tout $x$", il définit ainsi le quantificateur particulier ou existentiel $\exists x$ (ou $\exists y$ etc.), qui se lit "Pour quelque $x$" (ou "Il existe au moins un $x$ tel que") :

(2.27) $\exists x \Phi x \quad =_{def} \quad \neg \forall x \neg \Phi x.$

En utilisant le quantificateur, on peut dégager ainsi la forme des propositions 2.19 et 2.20.

La proposition

(2.28) Tous ceux qui intéressent tous les hommes intéressent Socrate

signifie

(2.29) Tout individu intéresse Socrate si tant est que cet individu intéresse tout être si cet être est homme.

Cette paraphrase mise en langue formelle ("est homme" devient $Hx$, "$x$ intéresse $y$" s'écrit $I(x, y)$ et "Socrate" se lit $s$) devient

(2.30) $\forall x\ [\forall y\ (Hx \supset Ixy) \supset Ixs]$,

tandis que

(2.31) Tous ceux que tous les hommes intéressent, intéressent Socrate

devient

(2.32) $\forall x\ [\forall y\ (Hy \supset Iyx) \supset Ixs]$.

Les propositions universelles affirmatives et négatives qui étaient des propositions *catégoriques* chez Aristote deviennent des propositions *hypothétiques* dans la logique de Frege : "Tous les $A$ sont $B$" devient "Pour tout individu, s'il est $A$, alors il est $B$", et "Aucun $A$ n'est $B$" devient "Pour tout individu, s'il est $A$, alors il est non $B$". A première vue, cela peut paraître contre-intuitif. A la réflexion, cette analyse logique est aussi acceptable que celle de Boole dont elle est la transcription dans le langage de la logique des propositions.

Dans la mise en forme booléenne des propositions universelles ("Tous les $A$ sont $B$" se lisant "La classe des $A$ est incluse dans la classe des $B$"), une proposition universelle dont le terme-sujet a comme extension la classe vide est toujours vraie, car la classe vide est incluse dans toutes les classes. La proposition

(2.33) Toutes les licornes sont blanches

est donc vraie.

Dans la mise en forme frégéenne, les quatre propositions fondamentales d'Aristote ($A$, $E$, $I$, $O$) reçoivent les mêmes conditions de vérité que chez Boole. Le fait que les propositions universelles aristotéliciennes $A$ et $E$ ne véhiculent aucune présupposition d'existence était rendu par le fait que les propositions de la forme "Tous les $A$ sont $B$" et "Tous les $A$ sont non $B$" sont comptées comme vraies quand l'extension de $A$ est la classe vide. Chez Frege, l'absence de présupposition d'existence attachée à $A$ est rendue par le conditionnel philonien. Une proposition affirmative universelle est une conditionnelle dans laquelle la conjonction "si ... alors ..." est prise dans le sens philonien. Elle est vraie si l'antécédent est faux. La proposition

(2.34) $\forall x$ ($x$ est licorne $\supset$ $x$ est blanche)

est vraie pour toutes les valeurs de $x$ qui rendent "$x$ est licorne" faux, c'est-à-dire pour tous les individus de l'univers du discours.

La même conception a été développée, indépendamment, par Ch. S. Peirce :

> Dire "Toute salamandre vit dans le feu" est simplement une manière de dire "Si n'importe quoi, $x$, est une salamandre, $x$ vit dans le feu". Ceci diffère d'un conditionnel — si tant est qu'il le fasse — seulement par

l'identification de l'$x$ que cette phrase comporte. En revanche il n'y a rien de conditionnel à dire : "Il y a une salamandre et elle vit dans le feu" (Peirce IV, 60, 41).

Le passage de la représentation booléenne à la représentation frégéenne et peircéenne des propositions fondamentales d'Aristote est aisée et naturelle. Bornons-nous à l'universelle :

(2.35) (a) Tous les hommes sont mortels.

(b) L'intersection de la classe des hommes et de la classe des non mortels est vide :
$\{x : x \text{ est homme}\} \cap \{x : \neg x \text{ est mortel}\} = 0.$

(c) La classe des êtres qui sont hommes et non mortels est vide :
$\{x : x \text{ est homme} \wedge \neg x \text{ est mortel}\} = 0.$

(d) Personne n'est membre de la classe des etres qui sont hommes et non mortels :
$\neg\exists x(x \in \{x \text{ est homme} \wedge \neg x \text{ est mortel}\}).$

(e) Personne n'est homme et non mortel :
$\neg\exists x(x \text{ est homme} \wedge \neg x \text{ est mortel}).$

(f) Pour tout $x$, $x$ est non homme ou $x$ est mortel :
$\forall x\ (\neg x \text{ est homme} \vee x \text{ est mortel}).$

(g) Pour tout $x$, si $x$ est homme, alors $x$ est mortel :
$\forall x\ (x \text{ est homme} \supset x \text{ est mortel}).$

### 2.1.7. *L'algèbre de Boole et l'idéographie de Frege*

La transcription en algèbre de Boole ne peut représenter la différence de sens (différence de conditions de vérité) entre les deux énoncés "Si tous signent la pétition, je serai surpris" et "De tous, s'ils signent la pétition, je serai surpris", mais elle a un autre défaut mentionné plus haut : elle peut formaliser les raisonnements de la logique des propositions et ceux de la logique des termes, mais au prix d'interprétations différentes et incompatibles. Elle sera donc incapable de représenter des raisonnements *mixtes* dont la validité repose sur l'exploitation *simultanée* des structures logiques de la logique des propositions (les connecteurs) et de celle de la logique des termes (les quantificateurs), comme le raisonnement 2.36 (adapté de Quine, 1973, p. 148).

(2.36) Si Pierre est digne de foi, alors Paul a fait du bon travail et aucun journaliste n'était absent,
Si Paul a fait du bon travail, alors quelques journalistes ou étaient absents ou rédigeaient,
donc si Pierre est digne de foi, quelques journalistes rédigeaient et n'étaient pas absents.

Dans la langue artificielle de Frege, *la mise en forme* de ces raisonnements mixtes ne présente aucune difficulté. C'est une première différence entre la formalisation booléenne et la formalisation frégéenne. On peut assigner au raisonnement ci-dessus la forme 2.37. Commençons par localiser les phrases atomiques :

| | |
|---|---|
| Pierre est digne de foi : | $p$ |
| Paul a fait du bon travail : | $q$ |
| $x$ est journaliste : | $Jx$ |
| $x$ est absent : | $Ax$ |
| $x$ rédige : | $Rx$ |

Le raisonnement 2.36 a la forme

(2.37) $$\begin{aligned} p &\supset [q \wedge \forall x(Jx \supset \neg Ax)] \\ q &\supset \exists x[Jx \wedge (Ax \vee Rx)] \\ p &\supset \exists x(Rx \wedge \neg Ax). \end{aligned}$$

En outre, et c'est là tout l'intérêt de cette mise en forme, il est possible de prouver la validité de ce raisonnement à l'aide des ressources déductives de l'idéographie de Frege.

Dans un texte que nous avons cité plus haut, Frege objectait que, chez Boole, entre deux parties de la logique, la logique consacrée aux rapports entre termes et la logique consacrée aux rapports entre propositions, il n'existe [...] aucun lien organique (Frege, cité par Cl. Imbert, 1971a, 25n). Dans la mise en forme de Frege, au contraire, et ceci est la deuxième différence, il existe bien un *lien organique* entre ces deux parties de la logique. Ce lien peut être mis en évidence à l'aide des équivalences 2.38 – 2.41 (où $\Phi$ est une formule de complexité quelconque dans laquelle $x$ apparaît libre au moins une fois) qui réduisent les propositions générales à des séquences finies ou infinies de propositions singulières (obtenues en substituant $a_1 \ldots a_n \ldots$ à $x$) :

(2.38) $$\forall x\Phi \equiv \Phi a_1 \wedge \cdots \wedge \Phi a_n \ldots$$

(2.39) $$\exists x\Phi \equiv \Phi a_1 \vee \cdots \vee \Phi a_n \ldots$$

(2.40) $$\forall x(\Phi \supset \Psi) \equiv (\Phi a_1 \supset \Psi a_1) \wedge \cdots \wedge (\Phi a_n \supset \Psi a_n) \ldots$$

(2.41) $$\exists x(\Phi \wedge \Psi) \equiv (\Phi a_1 \wedge \Psi a_1) \vee \cdots \vee (\Phi a_n \wedge \Psi a_n) \ldots$$

La troisième différence entre la formalisation de Boole et celle de Frege réside dans le fait que la première est un instrument de controle du raisonnement, un "*calculus ratiocinator*" au sens de Leibniz, qui permet de dégager la forme de

certains raisonnements (pas de tous), mais qui nous oblige à masquer le contenu. Comme Boole emploie le signe "×" dans un sens logique, il ne peut pas l'utiliser en même temps dans son sens arithmétique. S'il le faisait, il emploierait un symbolisme équivoque comme dans la proposition 2.42 qui affirme que l'intersection de la classe ayant comme membre le nombre 9 et de la classe ayant comme membre le nombre 2 est vide.

(2.42) $\{x : x = \sqrt{x} + \sqrt{x} + \sqrt{x} \neq \sqrt{x}\} \times \{x : x \times x = x + x\} = 0$

Ce symbolisme est équivoque car le signe "×" désigne tantôt l'intersection de classe (première occurrence), tantôt la multiplication entre nombres (deuxième occurrence). Or Frege assigne à son idéographie un double rôle : celui d'un *calculus ratiocinator* au sens de Leibniz, c'est-à-dire d'un calcul permettant de faire des déductions rigoureuses, mais aussi et en même temps celui d'une *lingua characteristica* permettant non seulement d'*exprimer* des concepts mathématiques — d'exposer la science acquise — mais aussi de *construire* de nouveaux concepts dont la définition servira dans les démonstrations (Frege, 1882, tr. fr. de Cl. Imbert, 1969, p. 212). L'exemple que Frege donne le plus souvent est celui de la définition de la limite d'une fonction (Cl. Imbert, 1969, p. 81) :

(2.43) $$\lim_{x \to x_0} f(x) = l \quad =_{def}$$
$$(\forall \epsilon > 0)(\exists \alpha > 0)(\forall x)(0 < |x - x_0| < \alpha \supset |f(x) - l| < \epsilon).$$

### 2.1.8. *L'univers du discours de la logique*

La langue naturelle ne contient pas de variables à proprement parler, mais elle renferme des signes qui se comportent comme des variables et qui peuvent être liés par des pronoms indéfinis qui sont l'équivalent des quantificateurs. Par exemple, la proposition

(2.44) Quelqu'un est entré et *il* m'a adressé la parole

contient le pronom anaphorique "il" qui joue le rôle d'une variable liée, rôle qui est mis en évidence par la mise en forme

(2.45) $(\exists x)(x$ est entré $\wedge$ $x$ m'a adressé la parole).

Ce qui différencie les variables quantifiées de la langue naturelle de celles de la langue formelle, c'est que la langue naturelle admet des domaines de référence distincts. Ceci a été souligné par R. Zuber dans *Structure présuppositionnelle du langage* (1972) :

> Chaque variable [de la langue naturelle] peut prendre des valeurs dans un univers qui est toujours restreint par une propriété présupposée : par exemple les indéfinis *quiconque*, *quelque chose*, *quelqu'un*, *n'importe qui* etc. ne sont pas absolument indéfinis, mais parcourent toujours un

> univers présupposé d'humains, ou d'objets physiques, etc. sans pouvoir parcourir la réunion ensembliste de ces ensembles (Zuber, 1972, p. 5).

La langue artificielle créée par Frege diffère ici encore de la langue naturelle. Dans "Logic as calculus and logic as language", J. van Heijenoort a bien mis en évidence cette différence :

> L'universalité de la logique s'exprime dans un trait important du système de Frege. Dans ce système, les quantificateurs qui lient des variables d'individus parcourent (*range over*) tous les objets [...] Boole a sa classe universelle, et De Morgan son univers du discours, dénotés par '1' [...] mais ceux-ci peuvent être changés à volonté. L'univers du discours comprend seulement ce que nous convenons de considérer à un moment donné, dans un certain contexte. Pour Frege au contraire, il ne peut être question de changer les univers. On ne peut même pas dire qu'il se restreint à *un* univers. Son univers, c'est *l'*univers. Pas nécessairement l'univers physique, évidemment, car pour Frege certains objets ne sont pas physiques (van Heijenoort, 1967b, p. 325).

Frege ne se prive pas pour autant des avantages des quantificateurs restreints qu'on trouve dans les exemples 2.46 et 2.47 :

(2.46) $\forall x \in \mathbf{N}(x \text{ est pair} \vee x \text{ est impair})$

(2.47) $\exists x \in \mathbf{N}(x \text{ est pair})$,

mais ces quantificateurs restreints sont des abréviations, introduites par les définitions suivantes :

(2.48) $\forall x \in \alpha,\ \Phi x =_{def} \forall x(x \in \alpha \supset \Phi x)$

(2.49) $\exists x \in \alpha,\ \Phi x =_{def} \exists x(x \in \alpha \wedge \Phi x)$.

Le champ de variation ou parcours de valeurs que Frege donne aux variables liées, en prenant comme il le fait l'univers comme univers du discours, a deux conséquences importantes. La première, c'est que pour lui, la logique n'est pas seulement un *calcul*, ni même seulement un *langage*, mais qu'elle est une *théorie* qui parle d'une réalité extérieure. Une position qui s'exprime dans l'aphorisme de Gonseth "la logique est la physique de l'objet quelconque", et avant lui chez Russell qui écrit dans son *Introduction to Mathematical Philosophy* (1919) :

> La logique s'occupe du monde réel tout autant que la zoologie, quoique ce soit de ses traits les plus généraux et les plus abstraits (Russell, 1919, p. 169).

La deuxième conséquence, c'est que la logique étant une langue universelle, on y est en quelque sorte enfermé. Comme l'écrit Hintikka dans "On the Development of the Model-theoretic Viewpoint in Logical Theory" :

> Selon les thèses de cette tradition, vous êtes (en quelque sorte) prisonnier de votre langage, vous ne pouvez en sortir, vous ne pouvez le ré-interpréter sur une grande échelle, et vous ne pouvez exprimer la sémantique du langage dans le langage lui-même (Hintikka, 1988, p. 1).

Il n'y a pas *un au-delà de la logique*, pas de *point de vue extérieur* d'où l'on pourrait poser la question de savoir si l'on a conduit à son terme la tâche de codification des inférences valides que l'on s'est assignée.

> La seule question de complétude qui puisse se poser, écrit J. van Heijenoort, c'est, pour employer une expression de Herbrand, une question *expérimentale* [...] Pouvons nous épuiser tous les modes intuitifs de raisonnement réellement utilisés en science ? (van Heijenoort, 1967b, p. 327).

## 2.2. Notions de syntaxe logique et de sémantique formelle

### 2.2.1. *L'avènement de la théorie des modèles*

Le concept de complétude qui vient d'être évoqué n'est pas pleinement satisfaisant. S'il est bien vrai que les logiciens ont "réellement réussi à reproduire, sous forme de démonstrations formelles[2], tous les raisonnements rigoureux qui avaient toujours été utilisés en mathématique" (Tarski, 1936, 1974, p. 144), on peut légitimement déplacer la question et se demander si les raisonnements utilisés par les mathématiciens suffisent à démontrer toutes les vérités des mathématiques.

Le logicien peut bien codifier de manière exhaustive la pratique des mathématiciens, il ne s'apercevra pas, ce faisant, des lacunes qui pourraient exister dans cette pratique. La conception de la logique comme système universel soutenue par Frege, Russell et Wittgenstein ne permet pas de conceptualiser cette éventualité, c'est-à-dire de concevoir la possibilité d'une telle lacune.

Dans le cas des sciences de la nature, les théories sont mises à l'épreuve de l'expérience. Une théorie physique doit pouvoir rendre compte de tous les phénomènes dont elle traite, ces phénomènes étant décrits dans des énoncés d'observation distincts des énoncés théoriques. Par exemple, la théorie de la gravitation n'a pu rendre compte des perturbations dans la trajectoire de Mercure, ce que la théorie de la relativité a réussi à faire. Dans le cas de la logique et des mathématiques, pour avoir l'équivalent de la dualité entre énoncés théoriques et énoncés d'observation, il faudrait que l'on puisse définir les notions de vérité logique ou de vérité mathématique *indépendamment* de la notion de théorème de logique et de théorème de mathématique. Une fois cette tâche accomplie, on pourrait alors définir la correction ou adéquation d'une théorie déductive et sa complétude d'une manière vraiment objective, sans faire référence à la pratique et à l'autorité des mathématiciens et des logiciens.

[2] Une démonstration formelle est une démonstration dont on peut dire si elle est correcte ou non *sans interpréter* les symboles qu'elle contient. Il suffit de s'assurer que le passage d'une ou de plusieurs lignes à une autre ligne obéit à certaines règles de manipulation de symboles.

Une théorie sera dite correcte ou adéquate si tous ses théorèmes sont vrais (et "vrai" ne signifie plus "démontrable") et elle est complète si toutes les vérités formulables dans cette théorie sont démontrables. Pour pouvoir poser la question de l'adéquation et de la complétude dans ce sens-ci, il faut, au préalable, définir de manière rigoureuse et indépendante les notions de théorème et de vérité. La première tâche a déjà été commencée par Frege et menée à son terme par Carnap dans *Logische Syntax der Sprache* (1934), la seconde le sera par Tarski.

Dans la logique contemporaine, on observe deux courants : l'un va de Frege à Quine en passant par Russell. Les représentants de ce courant de pensée voient dans la logique un langage (interprété) et même une théorie qui traite des aspects les plus abstraits de la réalité. L'autre va de Boole à Hintikka en passant par Skolem. Ses représentants voient dans la logique un calcul. Cette dernière conception de la logique permet de traiter les systèmes axiomatiques formalisés non comme des collections de formules *vraies absolument*, mais comme des collections de formules *vraies relativement à un modèle*, c'est-à-dire relativement à un choix de *domaine* pour les variables et d'*interprétation* pour les constantes.

### 2.2.2. *La distinction entre langage-objet et métalangage*

La dissociation entre la notion de vérité mathématique et celle d'axiome ou de théorème mathématique n'est pas nouvelle. Les inventeurs de la géométrie non euclidienne la maniaient parfaitement puisqu'ils se sont posé la question suivante : "l'axiome (on disait plutôt le postulat) des parallèles est-il vrai ?". Malheureusement, si on adopte la conception frégéenne de la logique comme langue et théorie universelle défendue dans la *Begriffsschrift*, on se prive de la possibilité de transposer à la logique la distinction entre axiome et vérité.

> On devrait s'aviser, écrit Russell en 1903, que la méthode qui consiste à supposer faux un axiome et à déduire ensuite les conséquences de cette supposition, méthode qui a été jugée admirable dans des cas comme celui de l'axiome des parallèles, n'est pas universellement applicable. Car tous nos axiomes [de logique] sont des principes de déduction; et s'ils sont vrais, les conséquences qui paraissent découler de l'adoption d'un principe opposé n'en découleront pas réellement, en sorte que les arguments que l'on tire de la supposition qu'un axiome est faux sont ici sujets à des sophismes (*fallacies*) d'un genre spécial (Russell, 1903, p. 15, cité par Goldfarb, 1979, p. 353).

La solution à cette difficulté, on la trouve dans la distinction entre langage-objet et métalangage dont Russell lui-meme a eu l'idée plus tard (voir son introduction au *Tractatus logico-philosophicus* de Wittgenstein, 1922). Nous empruntons à Kleene l'explication de cette dichotomie dont Carnap et Tarski feront un fécond usage :

> Nous voici devant un petit paradoxe : comment peut-on *traiter* la logique mathématiquement [...] sans utiliser la logique elle-même ? La solution de ce paradoxe est simple [...] Nous plaçons la logique que nous étudions dans une boîte et celle que nous utilisons dans cette étude dans une autre boîte. Au lieu de boîtes, nous pouvons parler de "langages". Quand nous étudions la logique, la logique que nous étudions appartiendra à une langue dite *langage-objet* parce que cette langue et la logique qui s'y trouve incluse sont l'objet de notre étude. Notre étude de cette langue et de sa logique, en y incluant notre emploi de la logique dans cette étude, nous la regardons comme appartenant à une autre langue, que nous appelons *langage de l'observateur*. Nous pouvons aussi parler de *logique-objet* et de *logique de l'observateur* (Kleene, 1972, p. 11).

Pour désigner le langage de l'observateur, on utilise communément le terme "métalangage".

Le *métalangage syntaxique* traite des propriétés (et relations) des expressions et formules en faisant abstraction de leur signification. Par exemple, les termes suivants appartiennent au métalangage syntaxique : "axiome", "théorème", "dérivation", "démonstration", "consistance syntaxique", "inconsistance syntaxique", "règle de transformation". En effet, pour savoir si une suite de propositions est un théorème, il ne faut examiner que les *transformations structurelles* par lesquelles cette proposition a été obtenue à partir des axiomes. Le terme relationnel "$Y$ est déductible de $X$" est également un prédicat du métalangage syntaxique.

Le *métalangage sémantique* traite des propriétés et relations des expressions qui reposent sur des rapports entre signifiant et signifié. Par exemple, les termes suivants appartiennent à la sémantique : "dénotation", "extension", "connotation", "intension", "validité", "satisfiabilité" (dans le sens de consistance sémantique), "inconsistance (sémantique)", "...est vrai de...", "...est une conséquence logique de...", "...est valide", "...implique...". Le prédicat métalinguistique "implique matériellement" exprime la *relation* qui existe entre deux propositions $A$ et $B$ lorsque le conditionnel philonien "Si $A$ alors $B$" est vrai. Par exemple,

(2.50) "la terre est plate" implique matériellement "$2 + 2 = 5$"

peut s'analyser en

(2.51) "Si la terre est plate, alors $2 + 2 = 5$" est vrai.

Pour définir l'adéquation et la complétude d'un sytème de logique, nous avons besoin de la notion de *conséquence logique*. Cette définition repose, à son tour, sur celle de *vérité dans un modèle*, laquelle repose sur la notion de *satisfaction* ("...est satisfait par..." est synonyme de "...est vrai *de*...").

Cette dernière notion sera définie au chapitre 7. Nous nous bornerons pour le moment à l'illustrer par des exemples : "x a conquis la Gaule" est *vrai de* ou *satisfait par* César; "$x$ a tué $y$" est *vrai de* ou *satisfait par* Brutus et César

pris dans cet ordre.

### 2.2.3. *Tableau récapitulatif*

| *Notions syntaxiques* | *Notions sémantiques* |
|---|---|
| • Une formule d'une théorie axiomatique formalisée est un *théorème* si elle résulte d'une démonstration à partir des axiomes. | • Une formule d'une théorie axiomatique formalisée est vraie dans un modèle si elle est satisfaite par toutes les suites d'individus membres de l'univers du discours[3]. |
| • Une formule A est *dérivable* d'un ensemble Γ de formules s'il existe une démonstration de A à partir des formules de Γ et des axiomes. | • Une formule A est une *conséquence logique* d'un ensemble Γ de formules si tous les modèles qui rendent vrais les membres de Γ rendent vraie A. |
| • Une théorie axiomatique formalisée est *complète*[4] si toute formule fermée ou sa négation est un théorème. | • Une théorie axiomatique formalisée est *complète*[4] si toute formule vraie est un théorème. |

## 2.3. Contraintes associées à la calculabilité

### 2.3.1. *La notion d'effectivité*

Dans son *Introduction to Mathematical Logic*, A. Church requiert d'un *système axiomatique formalisé* qu'il satisfasse aux exigences suivantes :

- la spécification des symboles primitifs doit être effective en ce sens qu'il doit y avoir une méthode par laquelle, chaque fois qu'un symbole est donné, on peut toujours déterminer de manière effective s'il figure ou non parmi les symboles primitifs;
- la définition d'une formule bien formée (*well formed formula*, abrégée en "wff") doit être effective en ce sens qu'il y a une méthode par

[3] Par exemple, $\langle 3, 5\rangle$ satisfait "$x < y$" mais $\langle 5, 3\rangle$ ne satisfait pas "$x < y$". Si on n'avait pas pris des *suites*, et si on avait écrit que 3 et 5 satisfont "$x < y$", on n'aurait pas pu rendre compte de cette différence. En effet, "et" est commutatif. Notons que les suites n'appartiennent pas au domaine. Elles sont simplement un outil du métalangage.

[4] Sur le sens de "complet", voir 7.7.2 et 7.7.4.

laquelle, chaque fois qu'une formule[5] nous est présentée, nous pouvons déterminer de manière effective si elle est bien formée ou non;

- la spécification des axiomes doit être effective en ce sens qu'il y a une méthode par laquelle, chaque fois qu'une formule nous est présentée, nous pouvons déterminer de manière effective si elle figure ou non parmi les axiomes;
- les règles d'inférence, prises ensemble, doivent être effectives dans le sens fort suivant : il doit y avoir une méthode par laquelle, chaque fois qu'une formule est présentée comme découlant d'autres formules par inférence immédiate, on puisse toujours déterminer, de manière effective, si l'inférence immédiate proposée est conforme ou non aux règles d'inférence.

De ces exigences il suit que la notion de preuve est effective en ce sens qu'il existe une méthode par laquelle, chaque fois qu'une suite finie de formules est donnée, on peut toujours déterminer de manière effective si celle-ci est une preuve ou non (Church, 1956, pp. 50–51).

Remarquons que ces exigences sont très naturelles. Avant de commencer la démonstration, on doit pouvoir dire quels sont les axiomes sur lesquels on s'appuie et quelles sont les règles auxquelles on se pliera. On ne peut, même dans un jeu, changer de règles en cours de partie sans risquer d'effacer toute distinction entre l'action de gagner ou de perdre la partie. Les exigences de Church sont aussi libérales qu'il est possible; en effet, on peut se donner un nombre infini d'axiomes, à condition qu'on puisse reconnaître en un nombre fini de pas si un candidat-axiome est recevable ou non. Il en va de même pour les règles d'inférence. Remarquons que cette exigence est satisfaite pour le système d'axiomes de l'arithmétique de Peano. Quand on se cantonne dans la logique du premier ordre, le système d'axiomes de Peano contient en fait une infinité d'axiomes. En effet, le dernier axiome de Peano, à savoir le principe d'induction mathématique qui s'énonce ainsi ($S$ étant le foncteur "le successeur immédiat de" :

$$(2.52) \quad (\Phi(0) \wedge \forall x(\Phi(x) \supset \Phi(S(x)))) \supset \forall x \Phi x$$

est en réalité un *schéma d'axiome* qui admet une infinité d'instances (toutes les instances que l'on obtient en remplaçant le marque-place $\Phi$ par une formule de la catégorie syntaxique requise. On peut, cependant, obtenir n'importe quelle instance de cet axiome en *une* étape : une substitution sur toutes les occurrences de $\Phi$.

Par exemple, prenant pour $\Phi x$ la formule $x = 0 + x$, nous obtenons

$$(2.53) \quad [0 = 0 + 0 \wedge \forall x(x = 0 + x \supset S(x) = 0 + S(x))] \supset \forall x(x = 0 + x).$$

[5] Nous appelons *formules* les assemblages bien formés de symboles.

Les exigences imposées par Church reviennent à dire qu'on doit pouvoir reconnaître *par algorithme* les formules, les axiomes, les règles et les démonstrations.

> Un *algorithme* peut être décrit comme une suite finie d'instructions, exprimées de manière précise et qui — lorsqu'on les confronte à une question d'une certaine espèce et qu'on les exécute à la lettre — se termine invariablement, tôt ou tard, par [la production de] la réponse correcte. Par exemple, la méthode enseignée à l'école primaire pour multiplier deux entiers peut être présentée comme un algorithme. Elle est conçue pour répondre aux questions d'une certaine espèce : "Quel est le produit de $m$ et $n$ ?" Ecrite dans une langue sèche et explicite, elle pourrait être exécutée mécaniquement (Lewis, Papadimitriou, 1980, p. 38).

### 2.3.2. *La notion de théorème est-elle effective ?*

La notion de théorème n'est pas nécessairement une notion effective. Pour que la notion de théorème soit effective, il faut qu'il existe un algorithme grâce auquel, pour n'importe quelle formule du système arbitrairement choisie, on puisse déterminer si celle-ci est un théorème ou non, c'est-à-dire qu'on doit disposer d'une méthode algorithmique pour fournir la preuve de la formule si celle-ci est un théorème et une méthode algorithmique pour établir qu'il n'existe pas de preuve si la formule n'est pas un théorème.

On appelle *problème de la décision pour la prouvabilité* la question suivante : "La notion de théorème est-elle effective pour le système formalisé S ?" Résoudre le problème de la décision, c'est exhiber un algorithme si la notion est effective, ou prouver qu'il ne peut exister un tel algorithme si la notion de théorème n'est pas effective. Le calcul des propositions est décidable. Il y a aussi des systèmes *semi-décidables* : les théorèmes sont reconnaissables par algorithme, mais les non-théorèmes ne le sont pas. Le calcul des prédicats est semi-décidable. On montrera en 2.4.5 l'intérêt des contraintes d'effectivité.

### 2.3.3. *De l'effectivité à la récursivité*

La notion d'*effectivité* utilisée dans les paragraphes qui précèdent est une notion intuitive aux contours flous. Pour démontrer que certaines notions ne sont pas effectives (comme la notion de "non-théorème" dans certains systèmes axiomatiques formalisés), on ne peut se contenter de cette notion floue. Il faudra donc substituer à la notion d'*effectivité* une notion techniquement précise. Une des manières possibles d'atteindre cet objectif est de remplacer la notion de "fonction effectivement calculable" par la notion de "fonction récursive générale". Une manière d'y parvenir est de substituer à la notion intuitive d'*effectivité* la notion rigoureusement définie de *récursivité*. La notion de

récursivité est empruntée à la théorie des nombres. L'incursion que nous allons faire dans cette partie des mathématiques n'est nullement une digression. Nous verrons plus loin que par la codification numérique des notions de syntaxe logique (formule, dérivation, théorème, etc.), la notion de récursivité jouera un rôle essentiel en théorie de la preuve.

Nous introduirons l'idée de récursivité en partant de la définition récursive de l'addition, définition qui fait usage de la fonction *successeur* ($S$) :

- $x + 0 = x$,
- $x + S(y) = S(x + y)$.

Cette définition est, en fait, un algorithme pour l'addition des nombres naturels, c'est-à-dire des membres de l'ensemble $\mathrm{N}_0 = \{ 0, S(0), SS(0), SSS(0) \ldots \}$.

Par des applications successives de cet algorithme et en faisant usage du principe de substituabilité des identiques de Leibniz (cf. section 1.5.2), on calcule aisément le résultat de l'addition

$$SSS(0) + SS(0)$$

au moyen du développement suivant :

$$(2.54) \quad \underbrace{SSS(0) + SS(0)}_{A} = S(\underbrace{SSS(0) + S(0)}_{B}),$$

$$(2.55) \quad \underbrace{SSS(0) + S(0)}_{B} = S(\underbrace{SSS(0) + 0}_{C}),$$

$$(2.56) \quad \underbrace{SSS(0) + 0}_{C} = \underbrace{SSS(0)}_{D}.$$

Dans le deuxième membre de l'équation 2.55, on remplace $C$ par $D$ :

$$(2.57) \quad \underbrace{SSS(0) + S(0)}_{B} = \underbrace{SSSS(0)}_{E}.$$

Dans le deuxième membre de l'équation 2.54, on remplace $B$ par $E$ :

$$(2.58) \quad \underbrace{SSS(0) + SS(0)}_{A} = S\,\underbrace{SSSS(0)}_{E}.$$

La définition récursive de l'addition est une instance du *schéma de récursion primitive* dans lequel $f, g$ et $h$ sont des marque-place auxquels doivent être substitués des foncteurs, c'est-à-dire des symboles de fonction :

$$(2.59) \quad \forall x_1, \ldots, x_n \in \mathrm{N}_0,\ h(x_1, \ldots, x_n, 0) = f(x_1, \ldots, x_n),$$

$$(2.60) \quad \forall x_1, \ldots, x_n, y \in \mathrm{N}_0,$$
$$h(x_1, \ldots, x_n, S(y)) = g(x_1, \ldots, x_n, y, h(x_1, \ldots, x_n, y)).$$

Ce schéma formalise bien la notion de *calcul fait pas à pas*. La fonction $h$

dépend d'un paramètre formé par un vecteur de variables $x_1, \ldots, x_n$ supposé fixé, et d'une variable $y$. Si l'on se donne $y = 0$, la deuxième équation permet de calculer la valeur de $h(x_1; \ldots, x_n, 1)$ à partir de la fonction $g$. Si on connaît la valeur de $h(x_1, \ldots, x_n, 1)$ pour $y = 1$, le schéma nous permet de calculer la valeur de $h(x_1, \ldots, x_n, y)$ pour $y = 2$ et ainsi de suite.

Il ne faut cependant pas présenter le schéma de récursion primitive comme n'étant rien d'autre que l'expression plus précise d'une idée intuitive. Comme le fait observer J. Ladrière :

> Nous n'avons pas besoin de nous appuyer sur la compréhension préalable que nous pourrions avoir de l'idée de récursion pour manier correctement une fonction récursive; il nous suffit de suivre les indications du schéma [...] En réalité, ce n'est pas l'idée intuitive qui explique le schéma, c'est celui-ci qui explique la notion de récursivité; c'est en maniant le schéma de récursion que l'on acquiert une compréhension exacte de ce que vise l'expression "calcul qui se fait pas à pas" (Ladrière, 1975, p. 254).

### 2.3.4. *Fonctions récursives primitives*

Alors que l'idée intuitive d'effectivité était impuissante à déterminer une frontière, la définition technique de récursivité y parvient et, par surcroît, fait apparaître l'existence de deux classes de fonctions récursives, celle des fonctions récursives primitives, et celle des fonctions récursives générales. La première est incluse dans la seconde.

On appelle *fonction récursive primitive*[6] un élément du plus petit ensemble de fonctions de $\mathbf{N}^n$ dans N qui satisfait les deux exigences suivantes :

1. Il contient la fonction $S(x)$, *successeur de*, la fonction *zéro*$(x)$ qui prend la valeur 0 quels que soient ses arguments et les *fonctions d'identification* $\mathrm{Id}_i^n(x_1, \ldots, x_i, \ldots, x_n) = x_i$ qui prennent pour valeurs l'un de leurs arguments.
2. Il est fermé[7] sur l'opération de composition de fonctions et sur l'opération de récursion primitive, opération décrite par le schéma de récursion primitive cité plus haut.

On appelle *fonctions récursives générales*[8] les fonctions générées comme les précédentes, mais avec l'aide d'une opération supplémentaire, à savoir la *minimalisation* qui produit une fonction à $n - 1$ places à partir d'une fonction

[6] Les fonctions récursives primitives furent introduites dès 1888 par Dedekind.

[7] Un ensemble est fermé pour une opération lorsque les résultats de cette opération sont tous à l'intérieur de l'ensemble considéré : l'ensemble des naturels est fermé pour l'addition, mais pas pour la soustraction ($2 - 5 = -3$).

[8] L'idée de fonction récursive générale a été conçue par Herbrand, mais sa définition exacte est due à Gödel (1934).

à $n$ places. La notion de fonction récursive générale sert à définir les notions d'*ensemble récursivement énumérable* et d'*ensemble récursif* grâce auxquelles nous pourrons reformuler les exigences de Church sans recourir à la notion *intuitive* d'effectivité (Mendelson, 1964, p. 121).

### 2.3.5. *Ensembles récursivement énumérables*

Un ensemble $X$ de nombres naturels est appelé *récursivement énumérable* si et seulement si (a) il est l'ensemble vide, ou (b) il est l'ensemble des valeurs (*range*) d'une fonction récursive générale unaire $f$, ce qui formellement s'écrit :

(2.61) $\quad n \in X \equiv \exists x(f(x) = n).$

Par exemple, l'ensemble des carrés est un ensemble récursivement énumérable :

(2.62) $\quad X = \{0^2, 1^2, 2^2, 3^2, 4^2 \ldots\}$

On dit de la fonction $f$ qu'elle *énumère* les membres de $X$ (Rogers, 1971, p. 218). Remarquons qu'elle n'énumère pas les non-membres, c'est-à-dire ici les nombres entiers naturels qui ne sont pas des carrés.

Un ensemble $X$ de nombres naturels est appelé *récursif* s'il est récursivement énumérable et si son complément est récursivement énumérable. En d'autres termes, un ensemble est récursif s'il existe une fonction récursive générale $f$ qui énumère les membres de $X$ et une autre, $f'$, qui énumère les membres du complément de $X$. On peut aussi définir l'ensemble *récursif* $X$ en disant que c'est un ensemble de nombres naturels dont la *fonction caractéristique* est une fonction récursive générale. Rappelons qu'on appelle fonction caractéristique d'un ensemble une fonction qui envoie sur **V**(le vrai) les membres de l'ensemble et sur **F**(le faux) les non-membres.

Considérons la relation binaire $R$ dont la fonction caractéristique est une fonction récursive générale et totale, c'est-à-dire une application (cf. 7.1). Cette relation binaire est un ensemble récursif de couples pris dans le produit cartésien $\mathbf{N_0} \times \mathbf{N_0}$. A partir de cette relation, formons l'ensemble $X$ de tous les membres du domaine qui sont reliés à pour le moins un membre du codomaine de la relation. Cet ensemble correspond à (est l'extension de) la fonction propositionnelle unaire "$(\exists y)R(n, y)$". On peut, en effet, écrire :

(2.63) $\quad n \in X \equiv (\exists y)R(n, y).$

On notera que, bien que la relation $R$ soit récursive, l'ensemble $X$ obtenu en liant le deuxième terme de la relation par un quantificateur est seulement récursivement énumérable. L'exemple suivant permet de voir pourquoi il en est ainsi. Prenons comme relation binaire la relation exprimée par la périphrase "$n$ est une voiture qui ne démarre que si on forme le mot de passe constitué par les chiffres du nombre $y$". Lions la variable "y" par un quantificateur existentiel. A l'aide de cette fonction propositionnelle on peut reconnaître les membres de $X$ et les énumérer, un à un, en essayant successivement les mots de passe codés

par les nombres pris en ordre croissant. Dès que pour un nombre la voiture démarre, on sait qu'elle appartient à l'ensemble $X$. En revanche, on ne peut pas reconnaître les non-membres. Si une voiture ne démarre pas pour un mot de passe, on ne peut en conclure qu'elle est membre du complément de $X$, c'est-à-dire qu'elle est membre de l'ensemble de voitures sans mot de passe, mais simplement en panne. Il en irait tout autrement si le quantificateur existentiel était borné. Par exemple, si l'on avait

(2.64) $\quad n \in X \equiv (\exists y < 13)R(n, y),$

on pourrait alors dire qu'une voiture est membre du complément de $X$ si elle *ne* démarre *pas* quand on forme comme mot de passe un nombre $i$ tel que $0 \leq i \leq 13$.

## 2.4. Théorie de la calculabilité et métalogique

### 2.4.1. *La codification numérique du langage*

Church exige que les ensembles de symboles, les ensembles de formules, les ensembles d'axiomes et ceux de règles soient tous des ensembles effectifs. On voudrait pouvoir remplacer la notion intuitive d'"ensemble effectif" par la notion technique d'"ensemble récursif", mais un obstacle nous arrête : les ensembles récursifs sont des ensembles de nombres naturels. Or les ensembles de symboles ou d'expressions ne sont pas des ensembles de nombres, mais des ensembles d'entités linguistiques. Gödel a levé cet obstacle en associant aux expressions, des nombres choisis de manière telle qu'il soit possible d'encoder le langage dans les nombres et de le décoder algorithmiquement. Exposer ici les détails de cette codification nous entraînerait trop loin. Disons simplement que chaque symbole et chaque expression, chaque axiome, chaque règle d'inférence et chaque démonstration possède un nombre-code qui lui est attribué grace à une méthode qui permet, par la suite, de recouvrer l'expression linguistique à partir du nombre-code. On a donné à ces nombres-codes le nom de "nombres gödéliens".

Dans un système axiomatique formalisé, une preuve est une suite finie de formules. Chaque membre de la suite est un axiome ou une formule obtenue à partir d'une ou de plusieurs formules antérieures par l'application d'une règle d'inférence. Maintenant que les formules, les axiomes et les règles sont codés par des nombres gödéliens, une preuve devient une séquence de nombres. Chaque membre de cette suite est soit le nombre gödélien d'un axiome, soit le nombre gödélien d'une formule inférée par une règle à partir d'une ou de deux lignes dénotées, à leur tour, par un nombre gödélien. La suite finie qui constitue la preuve est elle-même codée par un nombre. A partir du prédicat unaire "$y$ est le nombre gödélien de la preuve", on définit le prédicat binaire "$n$ est le nombre gödélien d'une proposition qui a pour preuve la démonstration codée par le nombre $y$".

On démontre que si l'ensemble des formules bien formées, l'ensemble des axiomes et l'ensemble des règles — ou plus exactement l'ensemble des nombres gödeliens qui correspondent aux éléments de ces ensembles — sont des ensembles récursifs, alors la relation binaire exprimée par la fonction propositionnelle "$y$ est le nombre de la preuve de $n$" est aussi une relation récursive (voir J. Shoenfield, 1967). Faisons un pas de plus et définissons "$n$ est un théorème" à partir de "$y$ est un nombre gödelien d'une preuve". "*Thm*" désigne l'ensemble des théorèmes et "*Pr*" est mis pour "$n$ a pour preuve $y$". On écrira donc :

(2.65) $\quad n \in Thm \equiv (\exists y) Pr(n, y).$

Etre un théorème, c'est être la dernière ligne d'*au moins* une preuve. Comme le quantificateur n'est pas borné, l'ensemble des théorèmes dérivables d'un ensemble récursif d'axiomes n'est lui-même pas récursif, il est seulement *récursivement énumérable* :

> Les séquences $S_i = \langle A_{i_1}, A_{i_2}, \ldots, A_{i_m} \rangle$, qui sont des preuves formelles à partir des axiomes $H$, peuvent être énumérées de manière effective. On peut le faire, par exemple, en énumérant *toutes* les séquences (mots) sur l'alphabet du langage $L$ et en biffant les séquences qui ne sont pas des preuves (Rabin, 1977, p. 601).

Ce faisant, on n'énumère pas les non-théorèmes. En effet, il n'y a pas de limite assignable à la longueur des preuves; tout ce qu'on exige, c'est qu'elles soient finies. Dès lors, si aucune des suites de 1 à 13 lignes n'est une preuve, on est en droit de dire qu'on trouvera peut-être une preuve parmi les suites de 14 lignes ou plus.

### 2.4.2. *Récursivité et complétude*

Si la théorie $T$ est *axiomatisable*[9] — en d'autres termes si l'ensemble de ses axiomes est récursivement énumérable — et si, par surcroît, pour toute formule $\phi$, on a $\vdash \phi$ ou $\vdash \neg\phi$, alors non seulement l'ensemble des théorèmes est récursivement énumérable, mais le complément de cet ensemble l'est aussi. Comme précédemment, on énumère mécaniquement toutes les suites de propositions et pour chacune de ces suites, on vérifie si la suite en question est une démonstration de $A_i$. Si elle ne l'est pas, on vérifie si elle n'est pas une démonstration de $\neg A_i$. On peut d'ailleurs faire faire le travail à deux machines travaillant en parallèle. Comme par hypothèse, $T$ est complète syntaxiquement, au bout d'un temps fini, l'une de ces deux propositions aura été pourvue d'une preuve. Lorsque la théorie est axiomatisable et complète, l'énumération

[9] Une théorie *axiomatisée* est une théorie dont l'ensemble des axiomes est récursif. Une théorie *axiomatisable* est une théorie dont l'ensemble des axiomes est récursivement énumérable.

des séquences de formules suivie des deux contrôles précités (sur $A$ et $\neg A$) constitue un algorithme de *décision* qui permet en un *temps fini* de dire, pour une formule $A$ arbitrairement choisie, si celle-ci est un théorème ou si elle n'en est pas un. Le résultat sera atteint en un temps fini parce que toute démonstration a une longueur finie.

On serait tenté d'objecter qu'en partant d'un axiome, on peut y ajouter une deuxième ligne d'une infinité de manières si bien qu'on s'épuisera à énumérer l'ensemble infini des suites formées de deux formules et qu'on n'arrivera jamais à celles qui sont faites de trois formules. On en déduit qu'il y a des démonstrations de longueur finie et des théorèmes que l'on n'atteindra jamais.

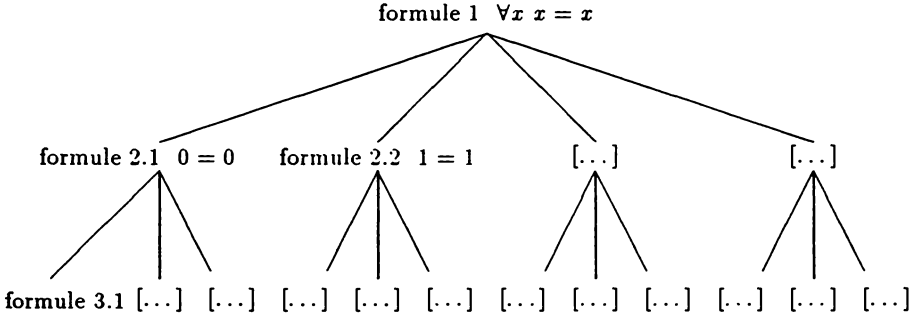



**Figure 2.1** *Balayage bidimensionnel de l'espace des suites*

Cette objection ne résiste pas à l'examen. Ce qu'on énumère, ce ne sont pas les suites de lignes en nombre croissant, mais les nombres; parmi ceux-ci, on repère, à l'aide d'un algorithme bien connu, les nombres gödéliens que l'on décode. On peut, à ce moment, reconnaître si l'on a affaire à un nom de symbole, de formule ou de suite de formules, et si la suite de formules est une démonstration ou non. Les nombres gödéliens sont ordonnés sur une seule dimension par la relation $<$. Rien n'empêche qu'une suite composée de trois formules ait un nombre gödélien plus petit qu'une suite composée de deux formules, mais où figurent des noms de grands nombres. On ne doit donc nullement avoir parcouru une suite infinie de dérivations de deux lignes avant d'examiner la première des dérivations à trois lignes.

Rappelons enfin qu'il ne faut pas confondre les deux énoncés 2.66 et 2.67; seul le premier est affirmé :

(2.66) Pour tout théorème $A$, il y a un temps fini à l'issue duquel $A$ aura été démontré par la méthode décrite par Rabin;

(2.67) Il y a un temps fini à l'issue duquel tout théorème $A$ aura été démontré par cette méthode.

Tout membre du complément de l'ensemble des théorèmes, c'est-à-dire tout

non-théorème, est identifié en un temps fini : en effet, dès que soit $A$, soit $\neg A$ est apparu comme dernière ligne d'une démonstration — ce qui se fait en un temps fini, comme on vient de le voir — la proposition opposée est déclarée être un non-théorème. Il en va tout autrement dans un système dans lequel il n'est pas vrai que toute formule fermée ou sa négation soit un théorème , comme par exemple le calcul pur des prédicats du premier ordre. Il y a dans ce dernier des formules telles que $(\exists x)Fx \supset (\forall x)Fx$. Comme ni cette formule ni sa négation n'est démontrable, on peut continuer indéfiniment à aligner les démonstrations à titre de preuves possibles, on n'arrivera pas en un temps fini à une réponse, c'est-à-dire à $\vdash (\exists x)Fx \supset (\forall x)Fx$ ou à $\vdash \neg(\exists xFx \supset \forall xFx)$.

On ne doit d'ailleurs pas s'en étonner : ni cette formule, ni sa négation ne sont valides. En effet, la formule en question est vraie dans un domaine à un seul membre et fausse dans les autres tandis que la négation de cette formule est fausse dans le premier cas et vraie dans les autres. Comme le calcul des prédicats est adéquat (correct), il ne permet de démontrer aucune des deux formules.

### 2.4.3. *Enumérabilité récursive et complétude*

Le calcul pur des prédicats du premier ordre (dépourvu de termes extra-logiques et d'axiomes extra-logiques) est *adéquat* — toutes les formules démontrables de ce calcul sont valides, c'est-à-dire vraies dans tous les modèles — et *complet* : toutes les formules valides sont démontrables. On a donc l'équivalence

(2.68) $\vdash A \equiv \models A$,

qui se lit

(2.69) "$A$" est un théorème si et seulement si "$A$" est valide.

L'ensemble des *formules démontrables* et l'ensemble des formules valides contiennent donc exactement les mêmes membres et dès lors coïncident.

Le calcul pur des prédicats du premier ordre satisfait les exigences d'effectivité de Church, donc l'ensemble des théorèmes de ce calcul est au moins récursivement énumérable. Comme l'ensemble des théorèmes de ce calcul et l'ensemble des *formules valides* n'en font qu'un, on conclut immédiatement que l'ensemble des formules valides est également récursivement énumérable et donc reconnaissable algorithmiquement en un nombre fini de pas.

Qu'en est-il des *formules non valides* ? En 1936, Church a démontré que le calcul des prédicats n'est que semi-décidable. Les théorèmes du calcul peuvent être énumérés récursivement, mais les non-théorèmes ne peuvent pas l'être. Il s'ensuit que l'ensemble des schémas non valides n'est pas récursivement énumérable. Il découle de là qu'on ne peut espérer trouver un algorithme pour prouver la consistance (la non-contradiction) d'un ensemble de formules du calcul pur des prédicats du premier ordre. Il existe, cependant, des algorithmes

pour prouver la consistance d'ensembles de formules de ce calcul d'un type particulier.

### 2.4.4. *Calculabilité, conséquence et déductibilité*

Nous avons vu au paragraphe 2.2.2 qu'une définition de la conséquence logique ou de la dérivabilité qui n'envisageait que les cas où les prémisses sont *vraies* était trop étroite et nous avons souligné les mérites d'une définition de la déductibilité qui permet de parler de propositions déductibles à partir de prémisses *fausses*.

On pourrait aller plus loin et adopter une conception plus libérale encore de la conséquence logique et de la déductibilité, une conception qui permette non seulement de parler d'inférences valides partant de prémisses fausses, mais aussi d'inférences valides à partir de prémisses *inconsistantes*[10]. Cet élargissement du concept de conséquence logique est-il opportun ?

Le théorème de semi-décidabilité du calcul des prédicats permet, rétrospectivement, de répondre à la question. Si on impose aux prémisses d'être consistantes, la relation de déductibilité dans le calcul des prédicats du premier ordre *cesse* d'être récursivement énumérable et effectivement axiomatisable, étant donné que l'ensemble des propositions consistantes n'est pas récursivement énumérable (van Benthem, 1985, p. 391). Nous avons donc une bonne raison technique de préférer la conception de la déductibilité la plus libérale, celle de Tarski, à celle de Bolzano qui exige que les prémisses soient consistantes.

### 2.4.5. *Décidabilité et complétude*

Le mathématicien, comme le note J. Bridge, ne s'intéresse pas aux formules $\Phi$ valides dans tous les modèles d'un système de logique comme le calcul pur des prédicats. Il s'intéresse plutôt aux formules $\Phi$ valides dans tous les modèles d'une théorie qui possède des termes extra-logiques et des axiomes propres telle que la théorie des groupes ou celle des corps commutatifs.

---

[10] Bolzano, qui a défini, cent ans avant Tarski, la notion sémantique de conséquence logique sous le nom de "dérivabilité", nom utilisé aujourd'hui pour désigner la relation syntaxique qui existe entre une formule qu'on peut dériver à partir d'autres formules par application des règles logiques, définit la relation de dérivabilité seulement pour des propositions compatibles entre elles : "Considérons le cas, écrit-il, où parmi les propositions compatibles $A, B, C, D, \ldots, M, N, O \ldots$ la relation suivante existe : toutes les idées, dont la substitution [des expressions] de ces dernières aux variables d'idées $i, j \ldots$ transforme une certaine partie de ces propositions, à savoir $A, B, C, D \ldots$, en vérités, ont la propriété de rendre vraies les propositions $M, N, O$ d'une autre partie de cet ensemble... J'appelle *dérivabilité* cette relation spéciale entre d'une part les propositions $A, B, C, D \ldots$ et d'autre part les propositions $M, N, O$" (Bolzano, 1837, 1972, p. 209).

Certaines théories mathématiques du premier ordre (formalisées dans le calcul des prédicats du premier ordre) sont décidables, en ce sens que pour toute proposition de ces théories, il est possible de décider par algorithme si elle est ou non un théorème. En d'autres termes, tant les non-théorèmes que les théorèmes seront repérables algorithmiquement. C'est le cas notamment de l'algèbre de Boole, de la théorie de l'ordre linéaire dense non borné et de l'arithmétique des nombres entiers limitée à l'addition. A ces théories qui ont un contenu mathématique relativement pauvre, viennent s'ajouter, notamment, certaines théories élémentaires de l'algèbre classique et de la géométrie traditionnelle dans lesquelles *toute formule fermée ou sa négation est un théorème* (la preuve est due à Tarski et McKinsey). Ces résultats sont importants dans la mesure où le contenu mathématique de ces théories est plus riche que celui des théories mathématiques mentionnées en premier lieu.

Si les ensembles des *théorèmes* de l'algèbre classique et de la géométrie traditionnelle sont récursifs, l'ensemble des formules vraies dans tous les modèles des axiomes propres de ces deux théories l'est aussi en vertu de l'équivalence "$\vdash A \equiv \models A$" et du fait que l'ensemble *Thm* en algèbre et en géométrie est récursif.

Cela signifie que pour toute formule valide de ces théories, il est possible — dans un temps fini et d'une manière mécanique — d'établir qu'elle est valide. Ce résultat est surprenant. Que l'on puisse identifier les théorèmes d'une théorie à l'aide d'un algorithme est moins étonnant. On peut, en effet, dire que l'inventeur d'un système d'axiomes et de règles d'inférence crée, en même temps que ces axiomes et ces règles, l'ensemble, éventuellement infini, de théorèmes qui en découlent, même s'il le fait de manière implicite. En revanche, la propriété pour une formule d'être valide dans tous les modèles qui vérifient les axiomes propres d'une théorie semble dépendre aussi peu de l'homme que la distribution des nombres premiers parmi les nombres naturels. On mesure ici l'importance des exigences d'effectivité formulées par Church en 2.3.1 Ces résultats surprenants n'ont été établis que pour les théories que l'on a axiomatisées en respectant les exigences de Church.

### 2.4.6. *Algorithmes intraitables et algorithmes performants*

L'affirmation que la procédure qui consiste à énumérer des suites de formules de plus en plus longues et à vérifier périodiquement si ces suites sont des preuves de "$A$" ou de "$\neg A$" est un algorithme paraît choquante :

> Cela est dû au fait qu'un trait essentiel d'un calcul algorithmique (*computation*) est que nous soyons sûr qu'il se terminera. Or quand quelqu'un donne l'*entrée* $A$ à l'algorithme, aucune borne à priori n'est donnée ici au nombre de pas avant que le calcul ne se termine (Rabin, Op. cit., p. 601).

A cette objection, Rabin répond que la complétude syntaxique de la théorie nous assure, en principe, que le calcul se terminera par la découverte d'une suite dont la dernière ligne est "$A$" ou dont la dernière ligne est "$\neg A$" et dont chaque ligne est tirée des précédentes par l'application d'une règle. La complétude syntaxique fait donc office de *preuve de terminaison* pour ce "programme".

Il y a des algorithmes totalement inefficients. Ce sont les algorithmes dont l'utilisation requiert trop de temps ou trop d'espace-mémoire :

> Un algorithme est considéré comme impraticable (*intractable*) s'il exige un nombre de pas ou une quantité d'espace qui croissent exponentiellement avec la longueur de l'*entrée* (Gallier, Op. cit., p. 50).

Il est clair que l'algorithme d'identification des théorèmes (ou des formules valides), reposant sur l'énumération récursive des preuves, appartient à cette espèce. Son intérêt théorique n'est cependant pas mis en péril par cette constatation. Comme l'écrit Wagner,

> nous inclinons à traiter comme accidentelle la question de savoir si certaines communautés peuvent comprendre des preuves de $10^{10}$ ou de $10^{100}$ ou de $10^{1000}$ étapes. Le raisonnement infinitaire [admettant des preuves de longueur infinie] semble plus radicalement impossible, et exclu par des aspects fondamentaux de la réalité [...] Les preuves finitaires semblent analogues à un genre naturel (Wagner, 1987, p. 31).

Nous verrons néanmoins que les logiciens qui travaillent dans le domaine de la démonstration automatique[11] sont préoccupés par le problème de la praticabilité, et nous présenterons des méthodes de preuve inspirées par le souci de restreindre l'*espace de recherche*.

[11] J. Siekman et G. Wrightson, 1983; Chin-Liang Chang et Richard-Tung Lee, 1973; P. Loveland, 1978; W. W. Bledsoe et D. V. Loveland (eds.), 1984; L. Wos, R. Overbeek, E. Lusk, J. Boyle, 1984 et Ph. Smets, E. H. Mandani, D. Dubois et H. Prade, 1988.

# Chapitre 3
# Le calcul des propositions – méthodes synthétiques –

## 3.1. Introduction

En logique, on distingue la *forme* du *contenu*. La logique des propositions est l'étude des raisonnements dont la forme est constituée par des variables propositionnelles ($p$, $q$, $r$...) et des connecteurs interpropositionnels tels que "et", "ou", "si...alors...", "non".

Une bonne logique des propositions doit nous permettre de *représenter* les raisonnements qui sont de son ressort, de distinguer ceux qui sont *corrects* de ceux qui ne le sont pas et d'expliquer où se trouve l'erreur lorsqu'ils sont fallacieux.

Avec un nombre fini de mots et de constructions grammaticales, on peut construire un nombre infini de phrases. De même, avec un nombre fini de propositions et de connecteurs interpropositionnels, on peut construire une infinité de raisonnements. Pour remplir son rôle, la logique des propositions doit élaborer des méthodes assez puissantes pour traiter n'importe lequel de ces raisonnements.

Comme on l'a vu dans la section 1.3.5, la distinction forme-contenu s'applique aux raisonnements.

**Définition 3.1 : le raisonnement.** Un raisonnement (*inférence*) est le passage de l'assertion d'une proposition ou d'un groupe de propositions (*prémisses*) à l'assertion d'une autre proposition (*conclusion*) (Hughes et Londey, 1965, p. 2).

**Définition 3.2 : la forme de raisonnement.** Une *forme de raisonnement*, ou *schéma d'inférence*, est une expression contenant des marque-place, telle que si ces marque-place sont remplis de manière appropriée, un raisonnement en résulte (Hughes et Londey, 1965, p. 3).

Le raisonnement suivant relève de la logique des propositions :

(3.1) Nous partons tous du "réalisme naïf", c'est-à-dire de la doctrine en vertu de laquelle les choses sont ce qu'elles paraissent être. Nous pensons que l'herbe est verte, que les pierres sont dures, et que la neige est froide. Mais la physique nous assure que la verdeur de l'herbe, la dureté des pierres et la froideur de la neige ne sont pas la verdeur, dureté et froideur que nous connaissons par notre propre expérience, mais quelque chose de différent... *Le réalisme naïf conduit à la physique et la physique, si elle est vraie, montre que le réalisme naïf est faux. Par conséquent, le réalisme naïf, s'il est vrai, est faux; par conséquent, il est faux* (Russell, 1940, pp. 23–24; l'italique est de nous).

La partie du raisonnement de Russell qui est en italique pourrait être réécrite ainsi :

(3.2) Si le réalisme naïf est vrai, la physique est vraie,
et si la physique est vraie, le réalisme naïf n'est pas vrai.
Donc si le réalisme naïf est vrai, le réalisme naïf n'est pas vrai.
Donc le réalisme naïf n'est pas vrai.

Dire "la proposition $p$ est vraie" équivaut à affirmer $p$. Nous pouvons dégager du *raisonnement* de Russell la forme suivante, sachant que "si...alors..." se note $\supset$, que "non" se note $\neg$ et que "et" se note $\wedge$ :

(3.3) $(p \supset q) \wedge (q \supset \neg p)$,
donc $p \supset \neg p$,
donc $\neg p$.

**Définition 3.3 : la proposition composée.** On appelle proposition composée toute proposition construite à partir d'une ou de deux propositions plus élémentaires à l'aide des connecteurs "non", "et", "ou", "si...alors...", "si et seulement si". Les trois propositions de 3.2 sont composées.

**Définition 3.4 : la proposition simple.** On appelle proposition simple une proposition qui n'est pas composée. Par exemple, l'énoncé

(3.4) le réalisme naïf est vrai

sera compté comme proposition simple selon ce critère. Pour dégager la forme d'un raisonnement, on remplace chaque proposition simple distincte par une variable propositionnelle distincte.

**Définition 3.5 : la forme propositionnelle.** La forme propositionnelle est à la proposition ce que la forme d'un raisonnement est à un raisonnement. On dégage la forme propositionnelle d'une proposition composée en remplaçant les propositions simples par des variables propositionnelles.

*Exemple.*

(3.5) $p \supset \neg p$

est la *forme propositionnelle* de la *proposition*

(3.6) Si le réalisme naïf est vrai, le réalisme naïf n'est pas vrai.

### 3.1.1. *Connecteurs naturels et connecteurs formels*

Le sens des connecteurs "non", "et", "ou", "si... alors...", "si et seulement si" en logique ne coïncide pas exactement avec le sens qu'ils ont dans la langue naturelle. Par exemple, en logique, le "et" est interprété de manière telle que 3.7 soit équivalent à 3.8 :

(3.7) Il a mis ses bas et [il a mis] ses chaussures.

(3.8) Il a mis ses chaussures et [il a mis] ses bas.

Dans la langue naturelle, ce n'est pas le cas. La conjonction "et" peut exprimer une idée de succession qui est absente du $\wedge$.

Il y a plusieurs manières de définir les connecteurs logiques : on peut les définir

> [...] implicitement et solidairement, par leur emploi dans un système d'axiomes,

mais on peut aussi les définir

> [...] explicitement et individuellement, par une table qui se réfère aux valeurs de vérité des variables propositionnelles sur lesquelles ils jouent (Blanché, 1970, p. 349).

Nous commencerons par présenter la logique des propositions sous la forme d'une combinatoire qui permet de vérifier la validité des raisonnements à l'aide de tableaux matriciels.

Dans un deuxième temps, nous présenterons la logique des propositions sous la forme d'un système axiomatique inspiré par le modèle euclidien qui a fasciné autant les philosophes — on songe ici à Spinoza — que les mathématiciens.

Nous présenterons ensuite la méthode de déduction naturelle, beaucoup plus souple et plus conforme à la manière dont nous raisonnons spontanément.

Nous examinerons dans le chapitre 4 des méthodes de preuve inspirées par le souci de limiter les *aléas* de l'invention et d'y substituer des procédés algorithmiques exécutables à la machine.

### 3.1.2. *Les fonctions de vérité*

Considérons les deux énoncés suivants :

(3.9) Les marées se produisent *parce que* s'exerce l'attraction universelle,

(3.10) Les marées se produisent *et* l'attraction universelle s'exerce.

Remplaçons dans 3.9 et dans 3.10 la phrase "l'attraction universelle s'exerce" par "la neige est blanche", c'est-à-dire par une phrase ayant une intension (signification) différente de la phrase initiale, mais ayant la même extension (la même valeur de vérité). Dans les deux cas, l'*intension* du tout est profondément altérée. Nous obtenons en effet deux nouveaux énoncés synthétiques :

(3.11) Les marées se produisent parce que la neige est blanche,

(3.12) Les marées se produisent et la neige est blanche.

L'effet du changement est cependant différent. Dans l'énoncé 3.11, l'intension et l'extension ont changé. Dans l'énoncé 3.12, seule l'intension a changé. L'extension est demeurée intacte, et elle serait demeurée intacte pour toute substitution à la phrase "l'attraction universelle s'exerce" (ou à la phrase "les marées se produisent") d'une phrase ayant même extension que la phrase que l'on remplace.

La conjonction "et" possède donc la propriété remarquable suivante qui fait défaut à la conjonction "parce que" : la valeur de vérité de la forme propositionnelle "$p$ et $q$" obtenue en infixant "et" entre deux variables propositionnelles varie en fonction de la valeur de vérité assignée aux variables propositionnelles $p$ et $q$. La conjonction "et" exprime donc une fonction qui prend ses arguments dans le produit cartésien $\{\mathbf{V},\mathbf{F}\} \times \{\mathbf{V},\mathbf{F}\}$ et ses valeurs dans $\{\mathbf{V},\mathbf{F}\}$. On l'appelle, pour cette raison, "fonction de vérité" et on dit que la conjonction "et" est vérifonctionnelle.

Les connecteurs "et" ($\wedge$), "ou" ($\vee$), "si... alors..." ($\supset$), "si... et seulement si..." ($\equiv$) et "non" ($\neg$) de la logique classique des propositions sont tous des foncteurs de vérité. On peut définir les fonctions qu'ils représentent par les matrices ou tables de vérité [1] de la figure 3.1.

### 3.1.3. *La bivalence*

La bivalence s'applique aux circuits électriques qui fonctionnent parfaitement. Les contacts monopolaires parfaits sont capables de *deux* valeurs et de deux seulement (P. Grosjean, 1970, p. 75).

La matrice du $\wedge$ décrit les interrupteurs en *série* (cf. figure 3.2). Le courant passe si et seulement si les contacts A et B sont fermés.

---

[1] Nous distinguons les *matrices* ou *tables de vérité* qui servent à définir les connecteurs, des *tableaux matriciels* qui sont des tables d'évaluation.

| $p$ | $\neg p$ |
|---|---|
| V | F |
| F | V |

| $\wedge$ | V | F | $q$ |
|---|---|---|---|
| V | V | F | |
| F | F | F | |
| $p$ | | | |

| $\vee$ | V | F | $q$ |
|---|---|---|---|
| V | V | V | |
| F | V | F | |
| $p$ | | | |

| $\supset$ | V | F | $q$ |
|---|---|---|---|
| V | V | F | |
| F | V | V | |
| $p$ | | | |

| $\equiv$ | V | F | $q$ |
|---|---|---|---|
| V | V | F | |
| F | F | V | |
| $p$ | | | |

**Figure 3.1** *Matrices (tables de vérité)*

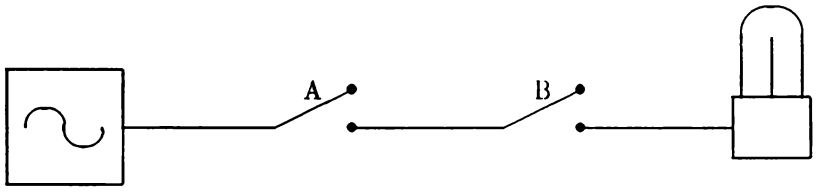



**Figure 3.2** *Interrupteurs en série*

La matrice du V décrit les interrupteurs en *parallèle* (cf. figure 3.3). Le

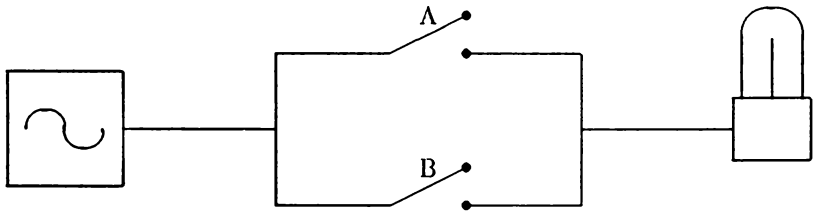



**Figure 3.3** *Interrupteurs en parallèle*

courant passe si et seulement si le contact A ou le contact B est fermé, ou les deux.

Le circuit du tiers exclu $p \vee \neg p$ se représentera par la figure 3.4.

Cette représentation des circuits permet de poser et de résoudre des problèmes pratiques tels que le suivant (Ayres, 1979, p. 242). On cherche à utiliser le plus petit nombre d'interrupteurs possible (cf. 9.4.6.4) :

(3.13) De son garage, $M$ peut pénétrer dans l'une des deux pièces. Trouver un circuit permettant à $M$ d'allumer ou d'éteindre la lumière soit du garage soit de l'une des deux pièces, quel que soit l'état des deux autres interrupteurs.

Peut-on étendre la bivalence aux énoncés de la langue naturelle, et affirmer que toute proposition est vraie ou fausse ? Aristote a soulevé l'objection des

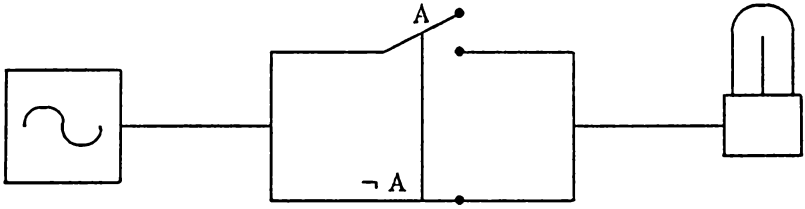



**Figure 3.4** *Circuit du tiers exclu*

futurs contingents, du type "Il y aura une bataille navale demain", qui semblent n'être ni vrais ni faux. Une des solutions proposées est d'admettre une logique à trois valeurs de vérité : le Vrai, le Faux, l'Indéterminé. Par ailleurs, G. Moisil, auteur des *Essais sur les logiques non Chrysipiennes* (1972), a montré qu'une logique *trivalente* inspirée de Lukasiewicz était nécessaire à la description et à l'analyse des dérangements dans les circuits électriques.

D'autre part, selon Belnap (*"How a Computer Should Think"* (1976/7) et *"A useful Four-Valued Logic"* (1976)), une machine sophistiquée capable d'un dialogue questions-réponses qui aurait la capacité de faire des inférences à partir de sa base de données devrait utiliser une logique *tétravalente* pour éviter que des incohérences mineures dans sa base de données ne conduisent, comme c'est le cas en logique classique, à des conclusions non pertinentes (en vertu du principe *ex falso sequitur quodlibet*).

Un type de propositions autre que les futurs contingents semble faire échec à la bivalence, ce sont les descriptions approximatives telles que "la France est hexagonale". Certains (Dubois et Prade, 1988) ont proposé d'adopter une logique floue, dans le fil des recherches de Zadeh. Sans mettre en question l'intérêt de ces développements profonds et originaux, nous pensons qu'il est possible de défendre le principe de bivalence en situant le *flou* et le *vague* sur une autre "dimension" que la valeur de vérité, suivant en cela Austin dans *Truth* :

> De nombreux autres adjectifs, dans la même classe que "vrai" et "faux", ont trait aux relations entre les mots [...] et le monde [...] Nous disons, par exemple, qu'un certain énoncé est exagéré ou vague ou hardi, d'une description nous disons qu'elle est grossière ou qu'elle induit en erreur [...] d'un compte rendu, nous disons qu'il est plutôt général ou trop concis. Dans des cas comme ceux-là, il n'est pas pertinent (*it is pointless*) d'exiger que l'on tranche en termes simples la question de savoir si l'énoncé est "vrai ou faux" [...] Est-il vrai que notre galaxie a la forme d'un œuf sur le plat, que Beethoven était un ivrogne, que Wellington a gagné la bataille de Waterloo ? Il y a des *degrés et dimensions* variés de succès dans [l'art de] faire des énoncés : les

énoncés s'accordent avec les faits d'une manière plus ou moins lâche, et de différentes manières à différentes occasions pour différents objectifs et des buts variés (Austin, 1970, pp. 129–130).

### 3.1.4. *La méthode des tableaux matriciels*

Une méthode fastidieuse, mais mécanique, permet de dire si une formule $X$ est vérifonctionnellement *valide* (c'est-à-dire valide en vertu de sa composition booléenne), *simplement consistante* ou *inconsistante* (il s'agit ici de la consistance et de l'inconsistance *sémantiques*, c'est-à-dire de la propriété, pour une formule, d'être vraie (respectivement fausse) pour quelque (respectivement aucune) interprétation. En anglais, le mot utilisé pour désigner la consistance sémantique est "*satisfiability*"). La formule $X$ est valide si, pour toutes les assignations possibles de valeurs de vérité à ses variables propositionnelles indépendantes, elle prend la valeur de vérité **V**. On calcule d'abord la valeur de vérité des fonctions booléennes qui ont comme arguments des variables propositionnelles, puis celle des fonctions booléennes qui ont comme arguments des fonctions et ainsi de suite jusqu'à ce qu'on arrive à l'évaluation de la formule elle-même. On procède de l'intérieur vers l'extérieur comme on le fait quand on veut calculer par exemple la valeur de l'expression $4^2+3!$, où l'on commence par calculer la valeur du carré de 4 et la factorielle de 3 et, après cela, la somme arithmétique des résultats des deux calculs précédents.

La figure 3.5 présente une évaluation et une vérification de la validité vérifonctionnelle d'une formule qui est une fonction de vérité à trois variables propositionnelles de la forme "$f(p, q, r)$" :

(3.14) $((p \wedge q) \supset \overline{r}) \equiv (p \supset (q \supset \overline{r}))$.

| $p$ | $q$ | $r$ | $\overline{r}$ | $p \wedge q$ | $q \supset \overline{r}$ | $(p \wedge q) \supset \overline{r}$ | $p \supset (q \supset \overline{r})$ | $[(p \wedge q) \supset \overline{r}] \equiv [p \supset (q \supset \overline{r})]$ |
|---|---|---|---|---|---|---|---|---|
| V | V | V | F | V | F | F | F | V |
| V | V | F | V | V | V | V | V | V |
| V | F | V | F | F | V | V | V | V |
| V | F | F | V | F | V | V | V | V |
| F | V | V | F | F | F | V | V | V |
| F | V | F | V | F | V | V | V | V |
| F | F | V | F | F | V | V | V | V |
| F | F | F | V | F | V | V | V | V |

**Figure 3.5** *Tableau matriciel d'une formule*

### 3.1.5. *Raisonnements et propositions*

Pour prouver la validité d'un raisonnement par la méthode des tableaux matriciels, on commence par le mettre sous forme implicative (on remplace le "donc" par "si... alors...", cf. section 1.4.4). Aux deux raisonnements 3.3 de Russell correspondent respectivement les formes propositionnelles valides

(3.15) $[(p \supset q) \wedge (q \supset \neg p)] \supset (p \supset \neg p)$,

(3.16) et $(p \supset \neg p) \supset \neg p$.

### 3.1.6. *Les inconvénients de la méthode des tables de vérité*

La méthode des *tables de vérité* nous oblige, sauf pour certaines classes de formules, à examiner indistinctement toutes les évaluations booléennes possibles. Or le nombre d'évaluations booléennes croît très rapidement : en logique bivalente, pour une formule comprenant $n$ variables propositionnelles distinctes, il y a $2^n$ évaluations booléennes. En logique bivalente, une interprétation d'une fonction de vérité $f(p_1, \ldots, p_n)$ est fixée dès qu'on sait quels $p_i$ ont la valeur **V** : ceux qui ne l'ont pas ont automatiquement la valeur **F**. La question "combien y a-t-il d'assignations positives distinctes possibles à $(p_1, \ldots, p_n)$" se ramène donc à la question "combien y a-t-il de sous-ensembles de $\{p_1, \ldots, p_n\}$". Or il existe une formule qui permet de calculer le nombre des *parties*, ou *sous-ensembles*, de l'ensemble $E$, c'est-à-dire le cardinal[2]de $\mathcal{P}(E)$ à partir du nombre de *membres* de $E$. C'est la formule

(3.17) Si $\#(E) = n$ alors $\#\mathcal{P}(E) = 2^n$.

Si l'ensemble $E$ comporte $n$ éléments, l'ensemble $\mathcal{P}(E)$ comporte $2^n$ éléments; pour cette raison, il est parfois noté $2^E$. On sait à présent que pour $n$ variables, il y a $2^n$ assignations possibles.

### 3.1.7. *L'algorithme de Quine*

Quine a proposé un algorithme qui rend plus maniable l'utilisation de la méthode des tables. On construit un arbre binaire de formules, orienté vers le bas, en prenant comme racine la formule à examiner et en substituant à chaque niveau **V** ou **F** aux variables propositionnelles, et en faisant les simplifications indiquées à la figure 3.6, où "$\rightarrow$" signifie "remplacer ce qui est à gauche de la flèche par ce qui est à droite".

Une formule est terminale quand elle contient uniquement des **V** ou des **F**. Il faut commencer par "ramifier" sur la variable qui fait le plus d'occurrences. Une

[2]Le signe # désigne le cardinal de l'ensemble, c'est-à-dire, pour les ensembles finis, le nombre de membres.

| | |
|---|---|
| $\neg \mathbf{F} \rightarrow \mathbf{V}$ | $\neg \mathbf{V} \rightarrow \mathbf{F}$ |
| $\mathbf{V} \supset X \rightarrow X$ | $X \supset \mathbf{V} \rightarrow \mathbf{V}$ |
| $\mathbf{F} \supset X \rightarrow \mathbf{V}$ | $X \supset \mathbf{F} \rightarrow \neg X$ |
| $\mathbf{V} \vee X$ ou $X \vee \mathbf{V} \rightarrow \mathbf{V}$ | $\mathbf{F} \vee X$ ou $X \vee \mathbf{F} \rightarrow X$ |
| $\mathbf{V} \wedge X$ ou $X \wedge \mathbf{V} \rightarrow X$ | $\mathbf{F} \wedge X$ ou $X \wedge \mathbf{F} \rightarrow \mathbf{F}$ |
| $\mathbf{V} \equiv X$ ou $X \equiv \mathbf{V} \rightarrow X$ | $\mathbf{F} \equiv X$ ou $X \equiv \mathbf{F} \rightarrow \neg X$ |

**Figure 3.6** *Algorithme de Quine : règles de simplification*

forme propositionnelle *valide* se reconnaît au fait que les feuilles de l'arbre sont toutes étiquetées **V**. Dans une forme propositionnelle *inconsistante*, elles sont toutes étiquetées **F**. Dans le cas où une feuille au moins est étiquetée **V**, mais où toutes ne le sont pas, la forme propositionnelle est *simplement* consistante, c'est-à-dire qu'un des sommets au moins de l'arbre doit être étiqueté **V**. Cet algorithme est plus rapide que la méthode des tables de vérité. L'application de l'algorithme au raisonnement 3.18 exposé dans la figure 3.7 requiert 23 lignes, alors que la table de vérité aurait nécessité 64 lignes ($2^6$).

## 3.2. Les méthodes déductives pour le calcul des propositions

Il existe aujourd'hui huit méthodes différentes de preuve, au moins, pour démontrer les lois du calcul classique des propositions :

- la méthode des tableaux matriciels (Peirce, Wittgenstein),
- la méthode axiomatique (Frege, Whitehead et Russell),
- la méthode de déduction naturelle (Gentzen, Jaskowski),
- la méthode des séquents (Gentzen),
- la méthode des tableaux sémantiques (Beth, Hintikka, Smullyan),
- la méthode des connexions et ses variantes (Bibel, Wallen),
- la méthode de réduction à la forme normale (Hilbert et Ackermann),
- la méthode de résolution (Robinson).

Si on adopte une perspective utilitaire et que l'on voit dans la logique une "boîte à outils", on a de bonnes raisons de les étudier toutes, car si elles ont toutes le même pouvoir démonstratif — elles permettent de démontrer les mêmes formules — il y a des raisons pratiques de préférer l'une à l'autre selon les cas à traiter.

Le point de vue utilitaire que nous venons d'évoquer est parfaitement légitime. Il peut se réclamer d'une tradition bien établie. Dès le deuxième siècle de notre ère, Alexandre d'Aphrodise donnait à l'ensemble des traités constituant la logique d'Aristote le nom générique d'*organon*, ce qui signifie "instrument".

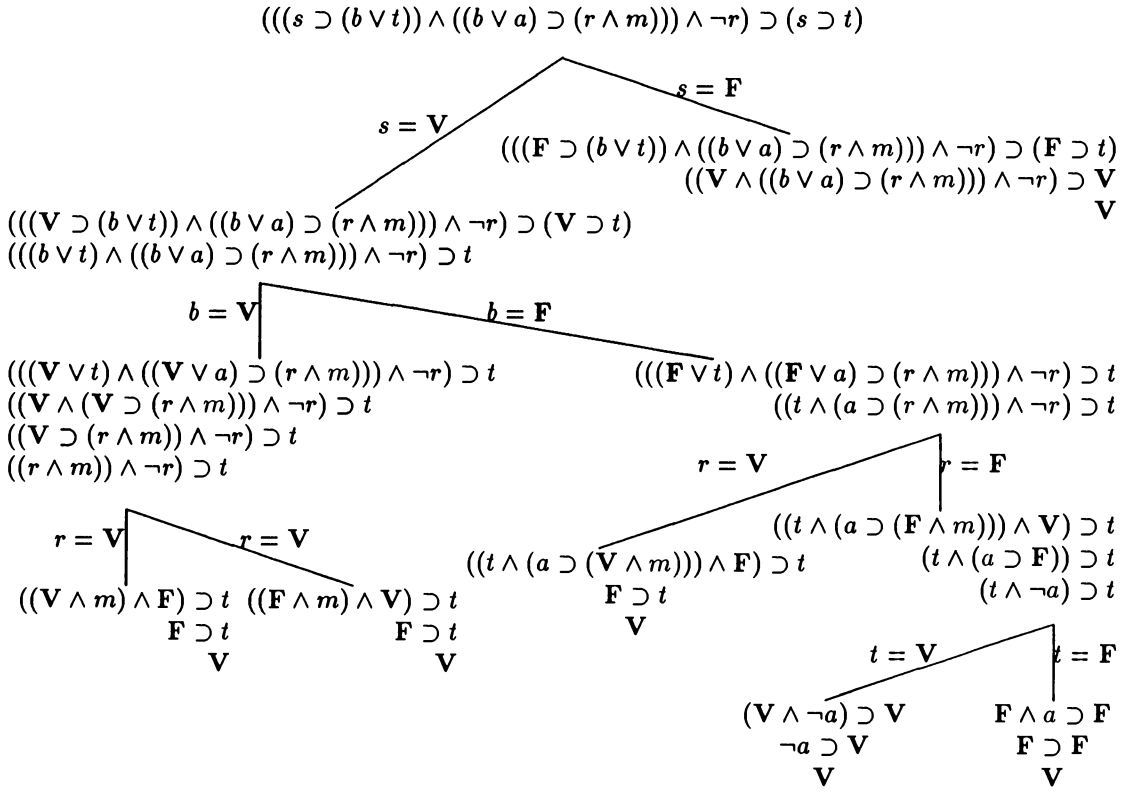



**Figure 3.7** *Algorithme de Quine : application*

Si l'on ne se borne pas à chercher dans la logique un instrument pour mieux raisonner, mais que l'on s'intéresse à la théorie de cet instrument, l'étude des huit méthodes est également nécessaire, mais il ne suffit plus d'exposer ces huit méthodes l'une à la suite de l'autre. Il faut les analyser et les comparer. Cette analyse et cette comparaison systématiques sont utiles au praticien. Elles accroissent en effet le rendement de l'instrument. Pour rendre la comparaison plus parlante, nous appliquerons les huit méthodes à la validation d'un même raisonnement, à savoir le raisonnement suivant (emprunté à I. Copi, 1967, et légèrement modifié; nous avons ajouté "s'il est venu seul" pour illustrer certaines règles) :

(3.18) S'il est venu seul, il a pris le bus ou le train. S'il a pris le bus ou son automobile, il est arrivé en retard et a manqué la réunion. Il n'est pas arrivé en retard. Donc, s'il est venu seul, il a pris le train.

## 3.3. Une langue formelle pour la logique des propositions

### 3.3.1. *La forme logique d'un raisonnement*

Le raisonnement 3.18 de la section précédente est formulé en langue naturelle. Pour dégager la forme logique de ce raisonnement, nous commençons par remplacer les connecteurs de la langue naturelle par les connecteurs correspondants de la langue formelle (correspondants, mais pas exactement synonymes, cf. 3.1), et par indiquer la portée des connecteurs à l'aide de parenthèses.

(3.19)
$I$ : il est venu seul $\supset$ (il a pris le bus $\vee$ [il a pris] le train),
$II$ : (il a pris le bus $\vee$ [il a pris] son automobile) $\supset$
(il est arrivé en retard $\wedge$ il a manqué la réunion),
$III$ : $\neg$(il est arrivé en retard),
*donc* : il est venu seul $\supset$ il a pris le train.

Nous avons dû répéter des bouts de phrases comme "il a pris" qui sont sous-entendus dans la langue naturelle. Cet effacement est utilisé par la langue naturelle pour indiquer que le connecteur "si ... alors ..." a la préséance sur "ou" dans $I$.

Dans un deuxième temps, on remplace chaque occurence d'une *même* proposition dépourvue de connecteurs par une occurence d'une variable propositionnelle, et chaque occurence d'une proposition *différente* dépourvue de connecteurs par une occurence d'une variable propositionnelle *différente*. Dans la pratique courante, on utilise comme variables propositionnelles les minuscules $p, q, r$ et pour disposer d'un stock infini $p', q', r', p'', q'', r'' \ldots$

Pour recouvrer plus aisément le raisonnement en langue naturelle à partir de sa forme, nous nous permettrons d'utiliser comme variables propositionnelles – contrairement à l'usage – toutes les lettres (minuscules) de l'alphabet et nous représenterons chaque proposition distincte par une lettre qui figure dans un mot de la phrase. Dans le raisonnement 3.18, ces lettres ont été soulignées.

La forme du raisonnement 3.18 sera donc :

(3.20) $I$ : $s \supset (b \vee t)$,
$II$ : $(b \vee a) \supset (r \wedge m)$,
$III$ : $\neg r$,
*donc* : $s \supset t$.

En parlant de "la forme logique" d'un raisonnement, nous donnons à penser qu'un raisonnement de la langue naturelle n'a qu'*une* forme. C'est inexact. A un même raisonnement de la langue naturelle, on peut souvent assigner des formes logiques différentes, si l'on dispose de langues formelles différentes.

Par exemple, selon qu'on dispose de la langue formelle du calcul des propositions ou d'une langue formelle plus riche, celle du calcul des prédicats, on pourra assigner au raisonnement 3.21 la forme logique 3.22 ou la forme logique 3.23.

(3.21) $I$ : Quelques bipèdes sont des mammifères,
$II$ : Tous les mammifères sont des vertébrés,
*donc* : Quelques bipèdes sont des vertébrés.

(3.22) $I$ : b,
$II$ : m,
*donc* : v.

(3.23) $I$ : $\exists x\ (Bx \wedge Mx)$,
$II$ : $\forall x\ (Mx \supset Vx)$,
*donc* : $\exists x\ (Bx \wedge Vx)$.

Pour arriver à 3.23, on a dû passer par une étape intermédiaire consistant à paraphraser la langue naturelle dans une langue semi-formelle :

(3.24) $I$ : $\exists x$ ($x$ est bipède $\wedge$ $x$ est mammifère),
$II$ : $\forall x$ ($x$ est mammifère $\supset$ $x$ est vertébré),
*donc* : $\exists x$ ($x$ est bipède $\wedge$ $x$ est vertébré).

Le choix entre les deux langues formelles (calcul des propositions et calcul des prédicats) n'est pas arbitraire. Seule la forme obtenue dans la deuxième langue permet de rendre compte du fait que le raisonnement en langue naturelle est valide. Il ne faudrait cependant pas conclure que lorsqu'on a le choix entre plusieurs "représentations formelles", il faut toujours préférer la plus riche. L'analyse optimale est l'analyse la plus grossière, qui soit cependant capable de rendre compte de la validité du raisonnement s'il est valide (Gochet, 1972, p. 26, d'après Copi 1967, Haack 1978, pp. 24–25). Par exemple, dans le raisonnement

(3.25) $I$ : Quiconque rendit visite à Pierre fut inquiété,
$II$ : Quiconque ayant parlé à Jeanne se souvenait d'elle,
$III$ : Personne ne se souvint de Jeanne,
*donc* : Jeanne ne rendit pas visite à Pierre.

on peut traiter "rendit visite à Pierre" comme un bloc d'un seul tenant, c'est-à-dire comme un prédicat à une place. Dès qu'une analyse grossière rend compte de la validité du raisonnement, toute analyse plus fine le fait aussi, mais est

superflue. Cette asymétrie doit être soulignée.

Si nous voulons utiliser la logique pour évaluer les raisonnements formulés en langue naturelle, nous devons commencer par embrigader ou enrégimenter, comme dit Quine, la langue naturelle dans une langue semi-formelle. Cela suppose un travail d'*encodage* et ensuite de *décodage* de la langue naturelle, dont il ne faut pas sous-estimer la difficulté.

Cette difficulté a été bien mise en lumière par P. Strawson dans son *Introduction to logical Theory* (1952) :

> Comparons ce que nous pourrions appeler "un syllogisme naturel" avec les exemples obstinément irréalistes que l'on trouve dans les livres. Un morceau de dialogue pourrait être :
>
> (3.26) *I* : Un homme vient d'avaler toute une bouteille d'alcool à brûler,
> *II* : Quiconque avale une telle dose ne survit pas,
> *donc* : Il mourra.
>
> Remarquez comment le premier terme indéfini conduit au pronom défini "il" de la conclusion, et comment le temps se modifie à chaque pas. A la différence des autres déviations par rapport au schéma strict (tel que "mourir" pour "ne pas survivre" et "avaler une telle dose" pour "boire toute une bouteille", etc.), ces deux-là ne peuvent être éliminés sans détruire le sens de l'original (Strawson, cité et traduit par F. Guenthner, 1987, p. 113).

Ceux qui, à juste titre, pensons-nous, rejettent la position malthusienne de Kemeny, qui a écrit que "le langage ordinaire ne convient pas aux arguments logiques (1957, p. 311)", doivent accepter de faire ce travail de paraphrase de la langue naturelle dans la langue formelle pour laquelle les notions de vérité, de conséquence logique et de dérivabilité sont rigoureusement définies. Le problème que pose cette paraphrase ou traduction de la langue naturelle dans la langue formelle, et inversement, ne sera pas traité ici[3].

Une difficulté supplémentaire réside dans le caractère incomplet (cf. 2.1.1) de la plupart des raisonnements en langue naturelle (Zuber 1990). Dans ces derniers, on sous-entend fréquemment des prémisses qu'il faut recouvrer pour rendre compte de leur validité.

### 3.3.2. *La langue formelle*

La langue formelle du calcul des propositions de Whitehead et Russell (1910) se présente de la manière suivante :

[3]Nous renvoyons le lecteur à l'ouvrage de J. Jayez, 1988, à celui de A. Thayse et al. 1989, à l'article de N. Cocchiarella, 1988.

1. Lexique de termes primitifs
   (a) un ensemble infini de variables propositionnelles : $\{p, q, r \ldots, p', q', r' \ldots, p'', q'', r'' \ldots\}$,
   (b) deux connecteurs primitifs : $\neg$, $\vee$.
2. Règles de formation
   (a) Toute variable propositionnelle est une *formule*.
   (b) Si $X$ est une formule, $\neg X$ est une formule.
   (c) Si $X$ et $Y$ sont des formules, $(X \vee Y)$ est une formule.
3. Règles de définition
   (a) $(X \supset Y) =_{def} (\neg X \vee Y)$.
   (b) $(X \wedge Y) =_{def} \neg(\neg X \vee \neg Y)$.
   (c) $(X \equiv Y) =_{def} ((X \supset Y) \wedge (Y \supset X))$.

Les règles de définition jouent un *double rôle* : elles font office de règle de formation et aussi — comme on le verra dans la section 3.4.8 — de règles d'inférence.

Dans leur rôle de règles de formation, elles peuvent être remplacées par l'adjonction des règles :

1. Si $X$ et $Y$ sont des formules, $(X \supset Y)$ est une formule.
2. Si $X$ et $Y$ sont des formules, $(Y \wedge Y)$ est une formule.
3. Si $X$ et $Y$ sont des formules, $(X \equiv Y)$ est une formule.

*Remarque.* Les variables $p$, $q$, $r$... et les connecteurs $\neg$, $\wedge$, $\vee$, $\supset$, $\equiv$ appartiennent au langage-objet. Les formules formées à l'aide des règles de formation appartiennent aussi au langage-objet. Par ailleurs, elles entrent dans la catégorie syntaxique des *formes propositionnelles*, c'est-à-dire dans la catégorie des phrases. Au contraire, les variables $X$, $Y$ et le symbole $=_{def}$ appartiennent au métalangage. Les métavariables $X$ et $Y$ appartiennent à la catégorie des *noms* (noms de formules plutôt que formules). On voit en effet que dans

(3.27) $X$ est une formule,

la variable occupe la position syntaxique d'un nom ou d'une locution nominale (d'un "syntagme nominal").

### 3.3.3. *La notion de constituant*

**Définition 3.6 : le constituant.** On appelle *constituant* ou *sous-formule* toute formule ou partie de formule qui est l'étiquette d'un nœud de l'arbre syntaxique de la formule.

*Exemple.* L'arbre syntaxique de
(3.28) $(\neg p \vee \neg(\neg q \equiv r))$
est décrit à la figure 3.8.

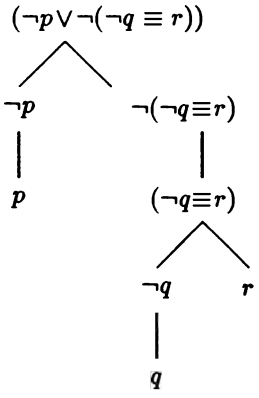



**Figure 3.8** *Arbre syntaxique d'une formule*

Les constituants sont donc $p, q, r, \neg p, \neg q, (\neg q \equiv r), \neg(\neg q \equiv r)$, et $(\neg p \vee \neg(\neg q \equiv r))$. Ne sont pas des constituants $\neg r$ ou $q \equiv r$.

La formule $(\neg p \vee \neg(\neg q \equiv r))$ est obtenue par les applications suivantes des règles de formation :

1. $p$ est une formule.
2. Si $p$ est une formule, $\neg p$ est une formule. Donc $\neg p$ est une formule.
3. $q$ est une formule.
4. Si $q$ est une formule, $\neg q$ est une formule. Donc $\neg q$ est une formule.
5. $r$ est une formule.
6. Si $\neg q$ et $r$ sont des formules, $\neg q \equiv r$ est une formule. Donc $\neg q \equiv r$ est une formule.
7. Si $\neg q \equiv r$ est une formule, $\neg(\neg q \equiv r)$ est une formule. Donc $\neg(\neg q \equiv r)$ est une formule.
8. Si $\neg p$ et $\neg(\neg q \equiv r)$ sont des formules, $(\neg p \vee \neg(\neg q \equiv r))$ est une formule. Donc $(\neg p \vee \neg(\neg q \equiv r))$ est une formule.

## 3.4. La méthode axiomatique

### 3.4.1. *Axiomes et évidence*

Dans son court traité intitulé *De l'art de persuader*, Pascal formule en deux maximes l'idéal du savoir déductif tel qu'il était incarné en son temps par la

géométrie :

> Cette véritable méthode, qui formerait les démonstrations dans la plus haute excellence s'il était possible d'y arriver, consisterait en deux choses principales : l'une, de n'employer aucun terme dont on n'eût auparavant expliqué nettement le sens, l'autre, de n'avancer jamais aucune proposition qu'on ne démontrât par des vérités déjà connues, c'est-à-dire, en un mot, à définir tous les termes et à prouver toutes les propositions (Pascal, réed. 1963, p. 349).

On ne peut prouver toutes les propositions. La raison de cette impossibilité a été clairement rappelée par Tarski :

> Nous prouvons chaque proposition sur base d'autres propositions et nous prouvons ces autres propositions à partir d'autres propositions à nouveau et ainsi de suite. Si nous voulons éviter un cercle vicieux et une régression à l'infini, la procédure de preuve doit s'arrêter quelque part (Tarski, 1969, p. 294).

On doit donc, continue Tarski, procéder comme Pascal le recommande (Tarski fait explicitement référence à l'opuscule de Pascal) :

> [...] Chaque discipline commence avec une liste d'un petit nombre de propositions, appelées axiomes ou propositions primitives, qui semblent intuitivement évidents et qui sont tenus pour vrais sans autre justification.

Cette *limitation* de la méthode axiomatique n'a pas changé depuis l'origine (les *Eléments* d'Euclide, vers 300 avant notre ère). Ce qui, en revanche, a changé fondamentalement, c'est la conception de la preuve.

Jusqu'aux dernières années du dix-neuvième siècle,

> une preuve était une activité par laquelle on visait à se convaincre et à convaincre les autres d'une vérité en discussion [...] Aucune restriction n'était imposée aux preuves, si ce n'est qu'elles devaient être intuitivement convaincantes (Tarski, *Ibid.*, p. 294).

Malheureusement, ce qui paraît intuitivement convaincant à l'un ne paraît pas nécessairement intuitivement convaincant à l'autre. Il y a des raisonnements fallacieux (les sophismes) dont l'inventaire s'est enrichi au fil du temps. Comment savoir si la preuve que nous utilisons n'en contient pas ?

Leibniz, le premier, comme nous l'avons vu (1.5.1), a conçu l'idée de preuves formelles dans lesquelles le passage d'une proposition à l'autre ne prendrait en considération que la *forme* des symboles à l'exclusion de leur *sens*, comme cela se fait dans les *calculs* en mathématique.

Cette conception formelle de la preuve retentit sur l'idée qu'on se fait des axiomes. Dans le contexte d'une preuve formelle, on ne doit retenir des axiomes que leur *forme*. On doit donc *faire abstraction* de leur caractère de propositions évidentes ou de propositions tenues pour vraies. En tant que parties d'une preuve formelle, les axiomes n'ont d'autre privilège que celui d'être des

*propositions primitives.*

Ici encore Leibniz a été un précurseur en exigeant

> qu'on démontrât ou donnât le moyen de démontrer tous les axiomes qui ne sont point primitifs, sans distinguer l'opinion que les hommes en ont et sans se soucier s'ils y donnent leur consentement ou non (Leibniz, vol. 4, p. 67).

Ce n'est cependant qu'avec Frege, Whitehead, Russell et Hilbert que les premières preuves formelles sont apparues. Une preuve formelle exige la création préalable d'une *langue formelle*. En effet, si nous faisons abstraction du *sens* des symboles, nous nous privons de la possibilité de distinguer les expressions bien formées des expressions mal formées, et nous devons suppléer à cette carence en formulant des règles de formation (cf. 3.3.2), règles grâce auxquelles nous pourrons reconnaître les formules en inspectant exclusivement leur forme (*shape*). Une fois en possession d'une langue formelle, on peut définir ainsi la preuve formelle :

> Une preuve formelle d'une proposition donnée est une *suite* finie de propositions telle que
>
> 1. la première proposition de la suite est un axiome,
> 2. chacune des propositions qui suivent est soit un axiome, soit directement dérivable de quelques-unes des propositions qui la précèdent dans la suite, en vertu des règles de preuve (qui également ne prennent en considération que la forme des formules),
> 3. la dernière proposition de la suite est la proposition à prouver.
>
> (Tarski, *Ibid.*, p. 295).

Le progrès réalisé par la formalisation est considérable :

> Le recours à l'évidence intuitive a été considérablement restreint : les doutes concernant la vérité des théorèmes n'ont pas été entièrement éliminés, mais ont été réduits à des doutes possibles concernant la vérité des quelques propositions énumérées comme axiomes et concernant l'infaillibilité des quelques règles de preuve (Tarski, *Ibid.*, p. 296).

Il ne faut pas conclure de ce passage que Tarski remet en cause le statut de formules *non interprétées* que possèdent les axiomes dans une preuve formelle. Il nous rappelle simplement que la mise entre parenthèses du sens des termes et de l'évidence des propositions est temporaire. Une preuve formelle n'est pas un jeu gratuit :

> après tout, la preuve formelle, tout autant que les anciennes preuves intuitives, est une procédure qui vise à l'acquisition de nouvelles propositions vraies (Tarski, *Ibid.*, p. 296).

L'objectif visé par les preuves, qu'elles soient formelles ou non, impose des contraintes extraformelles aux axiomes et aux règles d'inférence. Seule une

formule *interprétée* peut être qualifiée de vraie (ou fausse), ou de valide (ou consistante ou inconsistante). Il y aura donc une *phase* d'interprétation de la langue formelle et du système axiomatique formalisé, mais il reste acquis que la construction de la preuve ne fait pas intervenir l'interprétation.

Si on utilise la méthode axiomatique non pour *fonder* ou justifier certaines affirmations, mais pour *décrire* une structure, d'autres exigences que celles qui ont été mentionnées doivent être satisfaites, faute de quoi l'axiomatisation risque de ne rien ajouter à l'énumération pure et simple des vérités (on reviendra sur ce point dans le théorème 9.21).

### 3.4.2. *Le calcul des propositions axiomatisé*

Dans un système axiomatique formalisé, les axiomes ne sont pas des *propositions* primitives, c'est-à-dire des phrases pouvues de sens, mais des *formules* primitives dont on pourra dériver d'autres *formules* (les théorèmes) en appliquant des règles d'inférence qui seront, selon l'expression de Carnap, des *règles de transformation*.

On peut résumer ainsi la situation :

1. Les règles de formation isolent dans l'ensemble des assemblages le sous-ensemble des expressions bien formées (*formules*).

2. L'ensemble des formules comporte lui-même un sous-ensemble intéressant : les formules qui sont des *thèses* (c'est-à-dire les formules qui sont soit des axiomes, soit des théorèmes). La *théorie*, c'est-à-dire l'ensemble des thèses, est donc un sous-ensemble propre de la langue formelle, c'est-à-dire de l'ensemble des formules, sauf dans les systèmes inconsistants[4].

3. Enfin, on isole un sous-ensemble de thèses appelées *axiomes*, à partir desquelles on peut dériver les autres thèses, appelées théorèmes, à l'aide des règles d'inférence.

Ni pour distinguer les expressions bien formées des expressions mal formées, ni pour distinguer les axiomes des théorèmes, il n'a été nécessaire de faire allusion à l'*interprétation*, c'est-à-dire à la *signification* des symboles et assemblages de symboles. Dans un système axiomatique formalisé, on peut donc reconnaître qu'une formule est un axiome, qu'une suite de formules est une preuve et qu'une formule est un théorème, sans savoir *de quoi parlent* les formules.

[4] C'est précisément là la définition la plus générale de l'inconsistance : un système est inconsistant si toutes les formules sont des théorèmes.

### 3.4.3. *Le calcul des propositions interprété*

Nous allons examiner des présentations axiomatiques *du calcul des propositions*. Pour que ce calcul devienne une logique, il faut lui donner une interprétation. L'interprétation la plus courante assigne aux variables propositionnelles $p$, $q$, $r$... un domaine de variation constitué par $\{\mathbf{V}, \mathbf{F}\}$ et interprète les connecteurs comme des symboles de fonction de $\{\mathbf{V}, \mathbf{F}\}$ dans $\{\mathbf{V}, \mathbf{F}\}$, ou de $\{\mathbf{V}, \mathbf{F}\} \times \{\mathbf{V}, \mathbf{F}\}$ dans $\{\mathbf{V}, \mathbf{F}\}$, selon qu'il s'agit de foncteurs unaires ou de foncteurs binaires.

On choisit comme axiomes des *formes propositionnelles valides*, c'est-à-dire des *tautologies*, et on se fixe deux objectifs : on veut qu'en appliquant de manière répétée les règles d'inférence aux axiomes ou aux théorèmes déjà dérivés,

1. on n'engendre rien d'autre *que* des tautologies. Si cette condition est satisfaite, le système axiomatique (les axiomes et les règles) est dit *adéquat* ou *correct*,
2. on engendre *toutes* les tautologies. Si cette condition est satisfaite, le système est dit *complet*.

Muni d'une telle interprétation, le *calcul* des propositions qui était, en fait, un calcul des formules devient une *logique* des propositions ou plus exactement une logique des *formes propositionnelles*. En effet, il permet de démontrer à titre de théorèmes toutes les lois logiques qui *correspondent* (cf. 1.4.4) à des formes de raisonnement vérifonctionnellement valides[5].

### 3.4.4. *Logique des propositions ou des valeurs de vérité ?*

Comme les variables propositionnelles prennent comme valeurs des *valeurs de vérité*, c'est-à-dire le vrai ou le faux, plutôt que des *propositions* entendues comme intensions des phrases, on pourrait penser que les différences d'*intension* sont ignorées purement et simplement par la logique des propositions interprétée en *extension*.

Strawson avance un argument pour prouver que même l'interprétation extensionnelle (en termes de valeurs de vérité) du calcul des propositions fait place à l'*intension* (ou *sens*) des énoncés. Si on négligeait implicitement la proposition exprimée par les phrases et qu'on ne retenait que les valeurs de vérité de ces dernières, écrit Strawson, on aboutirait à une conséquence absurde : la multiplicité infinie des variables propositionnelles deviendrait une manière prolixe d'exprimer ce qui pourrait s'exprimer plus simplement avec les deux constantes **V** et **F** :

[5] C'est-à-dire valides en vertu des connecteurs vérifonctionnels et de l'agencement des variables propositionnelles.

> Si l'identité reflétée par l'identité typographique n'était que cela (à savoir l'identité de valeurs de vérité), nous ferions mieux d'abandonner le symbolisme de $p \supset p$ en faveur de $\mathbf{V} \supset \mathbf{V}$ et $\mathbf{F} \supset \mathbf{F}$... (Strawson, 1967, pp. 216–217).

L'*affirmation* de Strawson est correcte : même la logique extensionnelle réussit à capter des différences d'intension (de sens), en plus de la différence d'extension, plus grossière, entre **V** et **F**.

Mais *l'exemple* de Strawson est mal choisi. Pour faire voir que la langue formelle du calcul des propositions peut exprimer des différences intensionnelles, il aurait dû choisir une forme propositionnelle contenant plus de deux variables propositionnelles différentes comme, par exemple, l'axiome 3.45 de la section 3.4.5 :

$$(3.45) \quad ((p \supset q) \wedge (q \supset r)) \supset (p \supset r).$$

L'*identité typographique* entre deux variables reflète l'*identité intensionnelle* entre deux phrases de la langue naturelle qu'on peut faire correspondre aux deux variables; la *différence typographique* au sein de la même forme propositionnelle reflète la *différence intensionnelle* entre les phrases de la langue naturelle qu'on peut leur faire correspondre. Par exemple, l'axiome 3.45 précité dégage la forme commune à toutes les propositions composées chacune de *trois* atomes propositionnels distincts $(p, q, r)$. La restriction "au sein de la même forme propositionnelle" est nécessaire. Il faut, en effet, noter que des variables différentes dans des formes propositionnelles distinctes peuvent n'être que des variantes notationnelles l'une de l'autre comme, par exemple, $(p \supset p)$ et $(q \supset q)$. Certes l'*identité typographique* exprime aussi l'*identité extensionnelle*. En effet, deux phrases qui ont le même sens ont la même valeur de vérité. Mais la réciproque n'est pas vraie. Deux phrases peuvent différer par le sens sans différer par la valeur de vérité. C'est le cas de "la neige est blanche" et "l'herbe est verte".

### 3.4.5. *Le système axiomatique de Hilbert et Ackermann*

La première présentation axiomatique de la logique des propositions remonte à Frege (1879). Vient ensuite celle de Whitehead et Russell ($1^{ère}$ éd. 1910-1913). On doit à Hilbert et Ackermann (1934) un système axiomatique comportant cinq ensembles de trois axiomes et deux règles d'inférence. Il n'y a pas de définition. Chaque triplet d'axiomes a pour but d'isoler la contribution de chaque connecteur au travail déductif. Comme, sur les quinze axiomes, douze contiennent deux connecteurs, le succès est mitigé. Il n'en demeure pas moins vrai que ce système mérite d'être regardé comme le plus clair et le plus naturel (Kneale et Kneale, 1961, p. 527).

(3.29) $\quad p \supset (q \supset p)$[6]

(3.30) $(p \supset (p \supset q)) \supset (p \supset q)$

(3.31) $(p \supset q) \supset ((q \supset r) \supset (p \supset r))$

(3.32) $(p \wedge q) \supset p$

(3.33) $(p \wedge q) \supset q$

(3.34) $(p \supset q) \supset ((p \supset r) \supset (p \supset (q \wedge r)))$

(3.35) $p \supset (p \vee q)$

(3.36) $q \supset (p \vee q)$

(3.37) $(p \supset q) \supset ((q \supset r) \supset ((p \vee q) \supset r))$

(3.38) $(p \equiv q) \supset (p \supset q)$

(3.39) $(p \equiv q) \supset (q \supset p)$

(3.40) $(p \supset q) \supset ((q \supset p) \supset (p \equiv q))$

(3.41) $(p \supset q) \supset (\neg q \supset \neg p)$

(3.42) $p \supset \neg\neg p$

(3.43) $\neg\neg p \supset p$

Dans (Hubien, 1977, pp. 225–227), on trouve une version modifiée du système d'axiomes 3.29–3.43 qui ne vise plus à *isoler* la contribution de chaque connecteur au travail déductif mais qui, en revanche, a le mérite d'être plus claire et plus naturelle encore que celui de Hilbert.

Les axiomes 3.29 et 3.30 sont remplacés par la loi d'exportation

(3.44) $((p \wedge q) \supset r) \supset (p \supset (q \supset r))$.

Les axiomes 3.31, 3.34, 3.37 et 3.40 sont remplacés par leur forme importée

(3.45) $((p \supset q) \wedge (q \supset r)) \supset (p \supset r)$,

(3.46) $((p \supset q) \wedge (p \supset r)) \supset (p \supset (q \wedge r))$,

(3.47) $((p \supset q) \wedge (q \supset r)) \supset ((p \vee q) \supset r)$,

(3.48) $((p \supset q) \wedge (q \supset p)) \supset (p \equiv q)$.

Pour pouvoir déduire des théorèmes de ces axiomes, il faut des règles d'inférence.

[6] Nous nous permettons d'effacer les parenthèses extérieures des formules.

### 3.4.6. *Règles d'inférence*

**Substitution** : Si $X$ est une thèse, c'est-à-dire un axiome ou un théorème, si $p_1, \ldots, p_n$ sont des variables propositionnelles et si $Y_1, \ldots, Y_n$ sont des expressions bien formées, pas nécessairement distinctes, alors si $X'$ est $X$, où $Y_1$ a été substitué à $p_1, \ldots, Y_n$ à $p_n$, $X'$ est aussi une thèse. Formellement, "$X$ est une thèse" s'écrit $\vdash X$. L'expression $\vdash$ s'appelle en anglais *turnstile*.

*Remarque.* Les substitutions doivent toujours être faites *simultanément*.

On ne peut pas établir que le système est *adéquat*, c'est-à-dire que toutes les thèses sont des tautologies, en vérifiant chaque thèse à l'aide d'un tableau matriciel, car les thèses sont en nombre infini, mais on peut vérifier le caractère tautologique des axiomes et montrer ensuite que les règles d'inférence propagent la validité. On construit alors facilement une preuve par induction mathématique, établissant que toutes les thèses sont des tautologies.

Si $X$ est une tautologie, ce qui s'écrit formellement $\models X$, et si $X'$ est obtenue à partir de $X$ grâce à des substitutions conformes à la règle, alors $\models X'$.

#### 3.4.6.1. *Correction de la règle de substitution*

Puisque la forme propositionnelle $X$ est une tautologie, elle est satisfaite, c'est-à-dire rendue vraie, par n'importe quelle assignation de valeur de vérité à $p_1, \ldots, p_n$. Si nous assignons à $Y_1, \ldots, Y_n$ les mêmes valeurs de vérité qu'à $p_1, \ldots, p_n$, la valeur de vérité de $X'$ sera celle de $X$ pour ces assignations. Comme $X$ est une tautologie, cette valeur de vérité sera **V**. Donc $X'$ aura toujours la valeur **V** et sera donc aussi une tautologie (Mendelson, 1965, p. 19).

Dans la logique des propositions, substituer c'est toujours substituer une formule à une variable propositionnelle dans une thèse. Pour que *la portée* de la règle de substitution — qui est une règle d'inférence — soit bien délimitée, il faut que la notion de formule soit une notion effective au sens défini dans la section 2.3.1. En d'autres termes, il faut que l'ensemble des formules soit récursif (cf. 2.3.5). On voit ici pourquoi la méthode axiomatique requiert, pour être rigoureuse, d'être appliquée à une langue formelle, ce qui n'est pas réalisé, par exemple, dans les *Eléments* d'Euclide, ni dans l'*Ethique* de Spinoza, bien que les thèses de cette dernière soient — selon son auteur — "démontrées suivant l'ordre géométrique".

Notons que lors de l'application des règles de substitution, il arrive souvent que le nombre de parenthèses augmente. Par exemple, la substitution $\sigma$ notée par

$$(p \vee q)/p,$$

ou par

$$\langle p, (p \vee q) \rangle,$$

appliquée à l'axiome 3.35, donne lieu à

(3.49) $(p \vee q) \supset ((p \vee q) \vee q)$.

**Détachement :** La règle du détachement n'est autre que la règle du *modus ponens* restreinte aux formules qui sont des *thèses.* C'est pourquoi $X$, $X \supset Y$ et $Y$ sont précédés du signe $\vdash$ qui est absent du *modus ponens* :

$$\frac{\vdash X \supset Y \quad \vdash X}{\vdash Y}$$

3.4.6.2. *Correction de la règle du modus ponens et du détachement*

Les règles du *modus ponens* et du *détachement* sont correctes si et seulement si elles transmettent la "valeur distinguée" des prémisses de forme $X$ et $X \supset Y$ à la conclusion $Y$. Dans le cas du *modus ponens*, la valeur distinguée, c'est le vrai (**V**). Un tableau matriciel à quatre rangées suffirait à établir la correction de ce schéma d'inférence.

Dans le cas du détachement, ce qui doit être transmis des prémisses à la conclusion, c'est la *validité.* En d'autres termes il faut prouver que si les prémisses de forme $X \supset Y$ et $X$ sont des tautologies, la conclusion $Y$ l'est aussi, ce qui se note ainsi :

$$\frac{\models X \supset Y \quad \models X}{\models Y}$$

Pour établir la correction de la règle de détachement, on recourt à nouveau à un tableau matriciel. Celui-ci comportera $2^n$ rangées si le nombre des variables propositionnelles distinctes figurant soit dans $X$, soit dans $Y$, soit dans les deux, est $n$.

Pour démontrer la correction, on peut raisonner par l'absurde comme suit : on suppose que les prémisses de la forme $X \supset Y$ et $X$ sont valides et que la conclusion $Y$ ne l'est pas. Si $Y$ n'est pas valide, elle est falsifiable, ce qui signifie que parmi les $2^n$ assignations possibles de valeurs de vérité aux $n$ variables propositionnelles mentionnées au début de ce paragraphe, il y en a au moins une qui rend fausse la forme propositionnelle qui occupe la position de $Y$. Or, de deux choses l'une : ou cette assignation $\mathcal{A}$ falsifie aussi la formule occupant la place de $X$, ou elle ne le fait pas. Si elle le fait, alors la deuxième prémisse n'est pas une tautologie ($\not\models X$), si elle ne le fait pas, la première prémisse n'est pas une tautologie ($\not\models X \supset Y$) puisqu'il y a une assignation rendant vrai l'antécédent et faux le conséquent.

En résumé, si $\not\models Y$, alors ou bien $\not\models X$ ou bien $\not\models X \supset Y$. Donc si $\models X$ et

$\models X \supset Y$, alors $\models Y$[7].

### 3.4.7. *Le fragment implicationnel du calcul des propositions*

Pour illustrer le jeu combiné des règles de substitution et de détachement, il n'est pas nécessaire de traiter du calcul des propositions comportant les connecteurs $\neg$, $\wedge$, $\vee$, $\supset$, $\equiv$. On peut se limiter à un *fragment* de ce calcul, et à savoir au fragment implicationnel qui ne comporte qu'un seul connecteur $\supset$. Ce fragment a été axiomatisé en 1930 par A. Tarski et P. Bernays. Notre rapide évocation de ce fragment servira de toile de fond à l'exposé systématique des autres méthodes déductives.

*Symboles primitifs*

1. $p, q, r \ldots$ (variables propositionnelles),
2. $\supset$ (connecteur binaire),
3. Parenthèses.

*Règle de formation*

Si $X$ est une formule et si $X \supset Y$ est une formule, alors $Y$ est une formule.

*Axiomes*

| | | | |
|---|---|---|---|
| (3.50) | A1 | $(p \supset q) \supset ((q \supset r) \supset (p \supset r))$ | *Syllogisme* |
| (3.51) | A2 | $((p \supset q) \supset p) \supset p$ | *Loi de Peirce* |
| (3.52) | A3 | $p \supset (q \supset p)$ | *Simplification* |

*Règles d'inférence*

Les règles de substitution et de détachement.

[7] Appelons $P$ l'énoncé métalinguistique $\models X$, appelons $Q$ l'énoncé $\models X \supset Y$ et $R$ l'énoncé $\models Y$. Le raisonnement que nous avons développé dans la phrase commençant par "En résumé" a la forme suivante : $\neg R \supset (\neg P \vee \neg Q)$ donc $(P \wedge Q) \supset R$. Ce raisonnement saute certaines étapes que le lecteur pourra reconstituer, à titre d'exercice, grâce à la liste d'équivalence de 3.4.9

**Théorème 3.1 :** $((q \supset p) \supset r) \supset (p \supset r)$.

*Méthode de démonstration*

A ce stade, on ne dispose que de trois axiomes et de deux règles. Le théorème 3.1 ne peut être tiré d'aucun des trois axiomes par substitution. Il faudra donc utiliser la règle de détachement, donc employer deux des trois axiomes.

On cherche à réaliser la configuration de la figure 3.9, où le signe := se lit "devient".

$X :=$ axiome,
$X \supset Y :=$ axiome,
$Y :=$ théorème à démontrer.

**Figure 3.9** *Schéma d'application du détachement*

Aucune paire d'axiomes pris sous leur forme canonique ne peut entrer dans cette configuration. La règle du détachement n'est donc pas non plus utilisable isolément. Il ne reste qu'une issue : utiliser *conjointement* les deux règles. On tentera donc de faire entrer dans la configuration requise par la *règle de détachement* non les axiomes sous leur forme initiale, mais les axiomes transformés par une *substitution*. La recherche des substitutions adéquates exige un effort d'imagination.

Dans le cas présent, on partira de l'axiome A1, dans lequel on substituera $q \supset p$ à $q$ (voir les substitutions sous les petites accolades à la figure 3.10).

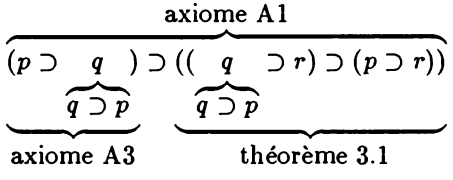



**Figure 3.10** *Démonstration du théorème 3.1*

La démonstration sous sa forme standard se présente ainsi :

| | | |
|---|---|---|
| (3.53) | $(p \supset q) \supset ((q \supset r) \supset (p \supset r))$ | Axiome 1 |
| (3.54) | $p \supset (q \supset p)$ | Axiome 3 |
| (3.55) | $(p \supset (q \supset p)) \supset (((q \supset p) \supset r) \supset (p \supset r))$ | subst. $q \supset p/q$ |
| (3.56) | $((q \supset p) \supset r) \supset (p \supset r)$ | 3.54, 3.55, dét. |

Au fur et à mesure que l'on démontre de nouveaux théorèmes, ceux-ci viennent s'ajouter aux axiomes comme formules à enchâsser dans la configuration

de *détachement*, éventuellement après transformation autorisée par la règle de *substitution*. La démonstration du théorème 3.2, par exemple, sera tirée d'une *instance* du théorème 3.1 (une "transformée" du théorème 3.1) et d'une *instance* de l'axiome A2 à l'aide d'une application de la règle du détachement. Alors que dans la démonstration du théorème 3.1, la substitution sur l'axiome A1 suffisait à faire apparaître l'axiome 3 dans l'antécédent de la transformée et à permettre un détachement, dans la démonstration du théorème 3.2, on devra opérer des substitutions sur les *deux* formules utilisées dans le détachement et pas seulement sur l'*une* d'entre elles.

**Théorème 3.2 :** $p \supset (p \supset (p \supset p))$.

Les substitutions à faire dans le théorème 3.1 et l'axiome A2 sont indiquées sous les petites accolades, à la figure 3.11.

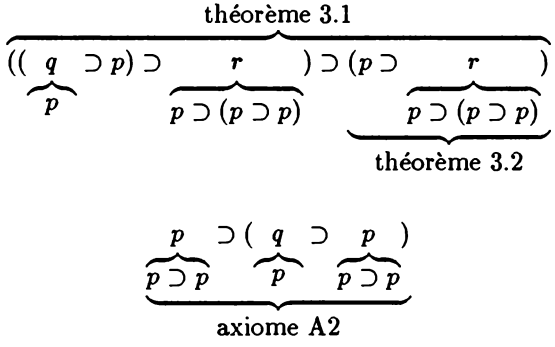



**Figure 3.11** *Démonstration du théorème 3.2*

*Remarques.* On voit que les substitutions dans une formule doivent être *simultanées* : si on avait substitué $p$ à $q$ dans l'axiome A3 et ensuite seulement $p \supset p$ à $p$, on aurait obtenu une instance de l'axiome A3 qui n'aurait pas coïncidé avec l'antécédent du théorème 3.1 et on n'aurait pas pu appliquer la règle de détachement.

A première vue, l'axiome A3 et l'antécédent du théorème 3.1 divergent par la structure : la position des parenthèses semble empêcher d'"assortir" l'axiome A3 et l'antécédent du théorème 3.1, mais un choix judicieux de substitutions écartera l'obstacle.

### 3.4.8. *Le calcul des propositions de Whitehead et Russell*

Le plus classique des systèmes axiomatiques du calcul des propositions est celui des *Principia mathematica* (1ère édition, 1910). Nous l'évoquerons

pour illustrer le fonctionnement des définitions qui sont absentes des systèmes évoqués plus haut. Le système de Whitehead et Russell a la même *puissance déductive* que le système d'axiomes présenté en 3.4.5. Les deux systèmes permettent de démontrer le même ensemble de théorèmes : ils sont dits, pour cela, *équipollents*.

(3.57) A1 $(p \vee p) \supset p$,

(3.58) A2 $q \supset (p \vee q)$,

(3.59) A3 $(p \vee q) \supset (q \vee p)$,

(3.60) A4 $(p \vee (q \vee r)) \supset (q \vee (p \vee r))$[8],

(3.61) A5 $(q \supset r) \supset ((p \vee q) \supset (p \vee r))$.

Dans le système de Whitehead et Russell, les symboles $\neg$ et $\vee$ sont pris comme primitifs. Les symboles $\supset, \wedge, \equiv$ reçoivent une définition *explicite*. Rien n'empêche de prendre ces derniers connecteurs aussi comme symboles primitifs. Les symboles primitifs[9] sont, dans ce cas, *implicitement* définis par les axiomes qui les contiennent. Les règles d'inférence sont le détachement et la substitution. Les définitions sont celles qui ont été exposées en 3.3. On notera une anomalie de ce système. Le connecteur $\supset$, qui est défini, apparaît dans les axiomes.

Les définitions jouent le rôle de règles d'inférence. La règle d'inférence qui leur est associée peut se formuler ainsi : si on remplace dans une thèse une formule par une autre qui lui est égale en vertu d'une définition, on retrouve une thèse.

**Théorème 3.3 :** $(q \supset r) \supset ((p \supset q) \supset (p \supset r))$.

*Démonstration.* Voir figure 3.12.

**Théorème 3.4 :** $(p \supset (q \supset r)) \supset (q \supset (p \supset r))$.

*Démonstration.* (analogue à la précédente).

**Théorème 3.5 :** $(p \supset q) \supset ((q \supset r) \supset (p \supset r))$.

---

[8] Bernays a démontré que l'axiome 4 n'est pas indépendant des autres. On peut le déduire de ces derniers par une longue chaîne de déductions. Pour prouver l'*indépendance* des axiomes A1, A2, A3, A5, on construit un modèle qui rend vrais A1, A2 et A3, mais faux A5, puis un modèle rendant vrais A1, A2, A5 et faux A3, et ainsi de suite (cf. 9.2.1 et 9.2.2.

[9] Ce système comporte comme *règles de formation* les règles 2b et 2c de la section 3.3.2. Les définitions 3a, 3b et 3c y jouent à la fois le rôle de règle de formation et de règle d'inférence. Dans ce dernier rôle, on fait appel au principe selon lequel le défini peut être remplacé par le définissant.

$(q \supset r) \supset ((p \vee q) \supset (p \vee r))$, axiome A5
$(q \supset r) \supset ((\neg p \vee q) \supset (\neg p \vee r))$ *Idem*, $\neg p/p$
$\neg p \vee q =_{def} p \supset q$
$\neg p \vee r =_{def} p \supset r$
$(q \supset r) \supset ((p \supset q) \supset (p \supset r))$

**Figure 3.12** *Démonstration du théorème 3.3*

*Démonstration.* Le théorème 3.5 s'obtient par détachement à partir des théorèmes 3.3 et 3.4. Les substitutions sont indiquées sous les petites accolades, à la figure 3.13.

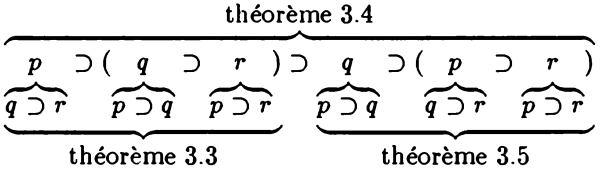



**Figure 3.13** *Démonstration du théorème 3.5*

### 3.4.9. *Métathéorème de l'échange*

Les démonstrations sont beaucoup plus faciles si l'on dispose de la règle d'inférence appelée *règle d'échange* ou *de remplacement*. Il s'agit d'une règle dérivée exprimée dans le métathéorème suivant :

(3.62) Si $\vdash A \equiv B$, si $C_A$ est une formule quelconque dont $A$ est un constituant et si $C_B$ diffère de $C_A$ en cela que $B$ a été mis à la place de $A$ dans une, plusieurs, ou toutes les occurrences de $A$ au sein de $C_A$, alors on peut déduire $\vdash C_A \equiv C_B$.

La règle d'échange diffère de la règle de substitution en cela qu'elle ne requiert *ni* que $C_A$ soit une thèse (axiome ou théorème), *ni* que le remplacement soit *uniforme*. En revanche, elle exige que les formules échangées satisfassent la condition $\vdash A \equiv B$.

*Remarque.* Les lettres $A$ et $B$ dans 3.62 ou dans la définition de la preuve à partir d'hypothèses (section 3.4.11) et dans le métathéorème de la déduction sont des *variables* du *métalangage*, au même titre que $X$ et $Y$ dans la règle de détachement.

*Démonstration.* On démontre cette règle à l'aide d'une induction mathématique. Cette induction requiert que l'on définisse d'abord la notion de *profondeur* (*depth*). Une formule $C_A$ est construite à partir de $A$ par des applications successives des clauses de la définition récursive de la notion de formule.

> Le nombre d'étapes dans cette construction de $C_A$ à partir de $A$, après que $A$ a été donné, et en excluant les étapes requises pour construire les parties ne contenant pas l'occurrence spécifiée de $A$, est ce que nous appelons la *profondeur* de l'occurrence de $A$ dans $C_A$. En d'autres termes, la profondeur de la partie $A$ dans $C_A$ est le nombre d'opérateurs dans la portée desquels il se trouve (Kleene, 1952, pp. 115–116).

*Induction.*

1. Si $A$ est de profondeur 0, alors $C_A = A$ et $C_B = B$, et l'affirmation que si $\vdash A \equiv B$ alors $\vdash C_A \equiv C_B$ se réduit à l'affirmation triviale que si $\vdash A \equiv B$ alors $\vdash A \equiv B$.

2. Si $A$ est de profondeur $d+1$, alors par hypothèse inductive, le théorème est supposé prouvé pour $d$. Considérons le cas où $C_A$ est de profondeur $d+1$. On examine successivement les cas suivants, où $M_A$ et $N$ sont des formules, $A$ étant à la profondeur $d$ dans $M_A$ :

| | | | |
|---|---|---|---|
| $C_A$ | est de la forme | | $\neg M_A$, |
| $C_A$ | " | " | $M_A \supset N$, |
| $C_A$ | " | " | $N \supset M_A$, |
| $C_A$ | " | " | $M_A \wedge N$, |
| $C_A$ | " | " | $N \wedge M_A$, |
| $C_A$ | " | " | $M_A \vee N$, |
| $C_A$ | " | " | $N \vee M_A$, |
| $C_A$ | " | " | $M_A \equiv N$, |
| $C_A$ | " | " | $N \equiv M_A$. |

3. Ensuite, on démontre neuf lemmes, correspondant à ces neuf cas. Le premier lemme s'énonce

$$\text{Si } \vdash M_A \equiv M_B \text{ alors } \vdash \neg M_A \equiv \neg M_B.$$

*Démonstration.* On commence par démontrer $\vdash (p \equiv q) \supset (\overline{p} \equiv \overline{q})$ dans le calcul des propositions[10].

[10] Pour nous faciliter la tâche, nous utilisons un système hybride : celui de Hilbert et Ackermann enrichi de la définition de $\equiv$, et nous utilisons dès à présent le métathéorème de la déduction qui sera exposé en 3.4.16. Quand la négation porte sur une proposition atomique ($p$, $q$, $r\ldots$), on peut la noter par un trait surmontant la variable propositionnelle ($\overline{p}$ est une variante pour $\neg p$).

| | | |
|---|---|---|
| 1 | $(p \wedge q) \supset p$ | Ax.3.32 de Hilbert et Ackermann |
| 2 | $((p \supset q) \wedge (q \supset p)) \supset (p \supset q)$ | $p \supset q/p,\ q \supset p/q$ |
| 3 | $(p \supset q) \supset (\overline{q} \supset \overline{p})$ | Ax.3.41 |
| 4 | $((p \supset q) \wedge (q \supset p)) \supset (\overline{q} \supset \overline{p})$ | Syllog. sur 2 et 3 |
| 5 | $(p \wedge q) \supset q$ | Ax. 3.33 |
| 6 | $((p \supset q) \wedge (q \supset p)) \supset (q \supset p)$ | $p \supset q/p,\ q \supset p/q$ |
| 7 | $(q \supset p) \supset (\overline{p} \supset \overline{q})$ | Ax. 3.41, $q/p,\ p/q$ |
| 8 | $((p \supset q) \wedge (q \supset p)) \supset (\overline{p} \supset \overline{q})$ | Syllogisme sur 6 et 7 |
| 9 | $((p \supset q) \wedge (p \supset r)) \supset (p \supset (q \wedge r))$ | Ax. 3.34 exporté |
| 10 | $((((p \supset q) \wedge (q \supset p)) \supset (\overline{p} \supset \overline{q})) \wedge$ $(((p \supset q) \wedge (q \supset p)) \supset (\overline{q} \supset \overline{p}))) \supset$ $(((p \supset q) \wedge (q \supset p)) \supset ((\overline{p} \supset \overline{q}) \wedge (\overline{q} \supset \overline{p})))$ | Ax. 3.34, $(p \supset q) \wedge (q \supset p)/p$ $\overline{p} \supset \overline{q}/q,\ \overline{q} \supset \overline{p}/r$ |
| 11 | $((p \supset q) \wedge (q \supset p)) \supset ((\overline{p} \supset \overline{q}) \wedge (\overline{q} \supset \overline{p}))$ | par détachement |
| 12 | $(p \equiv q) \supset (\overline{p} \equiv \overline{q})$ | par $X \equiv Y =_{def}$ $(X \supset Y) \wedge (Y \supset X)$ |

Ensuite, on substitue à $p$ et à $q$ les formules représentées par $\neg M_A$ et $\neg M_B$, de manière à obtenir

$(A_1)$ $\qquad \vdash (M_A \equiv M_B) \supset (\neg M_A \equiv \neg M_B).$

De là, et de $A_2$ qu'on a dans l'hypothèse du lemme

$(A_2)$ $\qquad \vdash (M_A \equiv M_B),$

il suit, par la règle de détachement, que

$(B)$ $\qquad \vdash (\neg M_A \equiv \neg M_B).$

En d'autres termes, $B$ est dérivable de $A_1$ et de $A_2$. Pour obtenir ce que l'on cherche, à savoir

(3.63) Si $\vdash (M_A \equiv M_B)$ alors $\vdash (\neg M_A \equiv \neg M_B)$,

il faut appliquer le principe suivant :

(3.64) Si $B$ est dérivable de $A_1, A_2, \ldots A_n$ alors le conditionnel "Si $A_n$ alors $B$" est dérivable de $A_1, A_2, \ldots A_{n-1}$.

Une formulation de ce principe dans laquelle "Si $A_n$ alors $B$" est remplacé par "$A_n \supset B$" sera donnée en 3.118 (voir aussi 3.4.16).

*Remarque.* La règle d'échange permet de remplacer un constituant $A$ d'une formule quelconque par n'importe quel constituant $B$ qui est logiquement équivalent à $A$. Nous avons vu qu'en particulier toute formule compte comme l'un de ses constituants. La règle d'échange nous autorise donc, pour les besoins de la démonstration, à remplacer même la prémisse entière $\mathcal{P}$ par une autre $\mathcal{P}'$ à condition que $\mathcal{P}'$ soit logiquement équivalent à $\mathcal{P}$.

Pour pouvoir exploiter pleinement la règle d'échange au sein d'un système axiomatique, il est utile de commencer par déduire des axiomes du calcul des propositions un ensemble de théorèmes d'équivalence qui servent souvent dans les démonstrations.

*Liste de théorèmes d'équivalence d'usage fréquent*

| | | | |
|---|---|---|---|
| (3.65) | **Identité** | (IDE) | $p \equiv p$, |
| (3.66) | **Lois de De Morgan** | (DM) | $\neg(p \wedge q) \equiv (\neg p \vee \neg q)$, |
| (3.67) | | | $\neg(p \vee q) \equiv (\neg p \wedge \neg q)$, |
| (3.68) | **Commutativité** | (COM) | $(p \wedge q) \equiv (q \wedge p)$, |
| (3.69) | | | $(p \vee q) \equiv (q \vee p)$, |
| (3.70) | **Associativité** | (ASS) | $(p \wedge (q \wedge r)) \equiv ((p \wedge q) \wedge r)$, |
| (3.71) | | | $(p \vee (q \vee r)) \equiv ((p \vee q) \vee r)$, |
| (3.72) | **Distributivité** | (DIS) | $(p \wedge (q \vee r)) \equiv ((p \wedge q) \vee (p \wedge r))$, |
| (3.73) | | | $(p \vee (q \wedge r)) \equiv ((p \vee q) \wedge (p \vee r))$, |
| (3.74) | **Contraposition** | (CONT) | $(p \supset q) \equiv (\neg q \supset \neg p)$, |
| (3.75) | **Implication matérielle** | (IM) | $(p \supset q) \equiv (\neg p \vee q) \equiv \neg(p \wedge \neg q)$[11], |
| (3.76) | **Equivalence matérielle** | (EM) | $(p \equiv q) \equiv (p \supset q) \wedge (q \supset p) \equiv (p \wedge q) \vee (\neg p \wedge \neg q)$, |
| (3.77) | **Idempotence** | (IDM) | $p \equiv (p \wedge p)$, |
| (3.78) | | | $p \equiv (p \vee p)$, |
| (3.79) | **Exportation-importation**[12] | (EX/IM) | $((p \wedge q) \supset r) \equiv (p \supset (q \supset r))$, |
| (3.80) | **Double négation** | (DN) | $p \equiv \neg\neg p$, |
| (3.81) | **Absorption** | (ABS) | $(p \wedge (p \vee q)) \equiv p$, |
| (3.82) | | | $(p \vee (p \wedge q)) \equiv p$. |

### 3.4.10. *Application*

Reprenons le raisonnement de Copi (p. 103, 3.18) sous la forme qu'il revêt dans *Symbolic Logic*, 1967 :

(3.83) Steve a pris le bus ou le train. S'il a pris le bus ou son automobile, il est arrivé en retard et a manqué la réunion. Il n'est pas arrivé en retard. Donc il a pris le train.

auquel nous pouvons assigner la forme logique

[11] L'expression 3.75 est une abréviation de trois biconditionnels. Il en va de même de 3.76

[12] On dit que $(p \supset (q \supset r)) \supset ((p \wedge q) \supset r)$ *importe* l'antécédent $q$ du conséquent dans l'antécédent et que $((p \wedge q) \supset r) \supset (p \supset (q \supset r))$ *exporte* $q$ dans le conséquent.

(3.84) $b \vee t$,
$(b \vee a) \supset (r \wedge m)$,
$\neg r$,
donc $t$.

Pour établir que ce raisonnement est logiquement valide à l'aide de la méthode axiomatique, on doit parvenir à déduire, à titre de théorème, la forme propositionnelle implicative qui lui correspond, c'est-à-dire une formule de la langue formelle. Dans le cas qui nous occupe, la formule à démontrer est la suivante :

(3.85) $((b \vee t) \wedge ((b \vee a) \supset (r \wedge m)) \wedge \neg r) \supset t.$

En d'autres termes, pour montrer qu'un raisonnement de la forme

(3.86) $X_1, \ldots, X_n \vdash Y$

est logiquement valide, nous cherchons à démontrer un théorème de la forme

(3.87) $\vdash X_1, \ldots, X_n \supset Y.$

Nous travaillerons dans le système de Hilbert et Ackermann en utilisant trois règles d'inférence (substitution, détachement, échange). On commence par démontrer un lemme appelé *polysyllogisme*.

**Théorème 3.6 :** $((p \supset q) \wedge (q \supset r) \wedge (r \supset s)) \supset (p \supset s)$.

*Démonstration.*

Forme (importée) du syllogisme :

$$\underbrace{(p \supset q) \wedge (q \supset r)}_{I} \supset \underbrace{(p \supset r)}_{II}$$

Forme (exportée) du syllogisme :

$$\underbrace{(p \supset r)}_{II} \supset \underbrace{((r \supset s) \supset (p \supset s))}_{III}$$

Par syllogisme, $((I \supset II) \wedge (II \supset III)) \supset (I \supset III)$, on obtient :

(3.88) $((p \supset q) \wedge (q \supset r)) \supset ((r \supset s) \supset (p \supset s))$,

et par application de la loi d'exportation-importation à la formule 3.88, on obtient le théorème 3.6.

Le polysyllogisme vaut pour $n$ antécédents (Lewis et Langford, 1932, 1959, p. 133). Nous utiliserons le théorème 3.7, dont la démonstration est laissée au lecteur.

**Théorème 3.7 :** $((p \supset q) \wedge (q \supset r) \wedge (r \supset s) \wedge (s \supset u)) \supset (p \supset u)$.

Dans le théorème 3.7, on fait les substitutions suivantes :

$$\begin{array}{lcl} \sigma_1 & : & \overline{r}/p, \\ \sigma_2 & : & \overline{r} \vee \overline{m}/q, \\ \sigma_3 & : & \overline{b} \wedge \overline{a}/r, \\ \sigma_4 & : & \overline{b}/s, \\ \sigma_5 & : & t/u. \end{array}$$

On obtient :

(3.89) $((\overline{r} \supset (\overline{r} \vee \overline{m})) \wedge ((\overline{r} \vee \overline{m}) \supset (\overline{b} \wedge \overline{a})) \wedge ((\overline{b} \wedge \overline{a}) \supset \overline{b}) \wedge (\overline{b} \supset t)) \supset (\overline{r} \supset t)$

et ensuite la forme exportée de 3.89 :

(3.90) $\underbrace{(\overline{r} \supset (\overline{r} \vee \overline{m}))}_{\text{axiome 3.35}} \supset$

$((((\overline{r} \vee \overline{m}) \supset (\overline{b} \wedge \overline{a})) \wedge ((\overline{b} \wedge \overline{a}) \supset \overline{b}) \wedge (\overline{b} \supset t)) \supset (\overline{r} \supset t))$ par dét.

On permute les deux premiers membres de la conjonction 3.4.10

(3.91) $(((\overline{b} \wedge \overline{a}) \supset \overline{b}) \wedge ((\overline{r} \vee \overline{m}) \supset (\overline{b} \wedge \overline{a})) \wedge (\overline{b} \supset t)) \supset (\overline{r} \supset t)$,

pour obtenir la forme exportée de 3.91 :

(3.92) $\underbrace{((\overline{b} \wedge \overline{a}) \supset \overline{b})}_{\text{axiome 3.32}} \supset (((\overline{r} \vee \overline{m}) \supset (\overline{b} \wedge \overline{a})) \wedge (\overline{b} \supset t)) \supset (\overline{r} \supset t)$

(3.93) $(((\overline{r} \vee \overline{m}) \supset (\overline{b} \wedge \overline{a})) \wedge (\overline{b} \supset t)) \supset (\overline{r} \supset t)$, par dét.

Par permutation dans l'antécédent et par contraposition appliquée à $(\overline{r} \vee \overline{m}) \supset (\overline{b} \wedge \overline{a})$, on obtient

(3.94) $((\overline{b} \supset t) \wedge (\neg(\overline{b} \wedge \overline{a}) \supset \neg(\overline{r} \vee \overline{m}))) \supset (\overline{r} \supset t)$.

Par application de la loi d'équivalence, on obtient

(3.95) $((b \vee t) \wedge ((b \vee a) \supset (r \wedge m))) \supset (\overline{r} \supset t)$.

Par importation sur 3.95, on obtient

(3.96) $((b \vee t) \wedge ((b \vee a) \supset (r \wedge m)) \wedge \overline{r}) \supset t$.

En pratique, pour trouver une démonstration de ce genre, on part du théorème à démontrer et on effectue des transformations sur base des lois d'équivalence jusqu'à ce qu'on obtienne une formule équivalente à celle qu'on veut démontrer, mais qui permet de lire en filigrane les thèses à prendre comme point de départ. Par exemple, en remontant du théorème à démontrer par des applications de la règle d'échange, on arrive à

(3.97) $(((\overline{r} \vee \overline{m}) \supset (\overline{b} \wedge \overline{a})) \wedge (\overline{b} \supset t)) \supset (\overline{r} \supset t)$.

On voit que grâce à une importation du $\overline{r}$ dans l'antécédent suivie d'une permutation qui l'amène en tête, on obtient

(3.98) $(\overline{r} \wedge ((\overline{r} \vee \overline{m}) \supset (\overline{b} \wedge \overline{a})) \wedge (\overline{b} \supset t)) \supset t$.

L'inspection de la formule fait apparaître en creux les chaînons manquants, à savoir les instances suivantes des axiomes :

(3.99) $\overline{r} \supset (\overline{r} \vee \overline{m})$,

(3.100) $(\overline{b} \wedge \overline{a}) \supset \overline{b}$.

La démonstration est longue. On voudrait pouvoir éviter le passage de 3.84 à 3.85, par lequel on transforme la forme de raisonnement :

(3.101) $X_1, \ldots, X_n$ donc $Y$, où $X$ et $Y$ sont des variables de formules

en forme implicative du type :

(3.102) $X_1, \ldots, X_n \supset Y$.

Pour éviter cette étape, il faudrait disposer d'un système axiomatique autorisant des preuves dans lesquelles peuvent figurer non seulement des thèses du système (axiomes ou théorèmes) qui sont toutes des lois logiques, mais aussi des prémisses quelconques, appelées "hypothèses".

### 3.4.11. *La preuve à partir d'hypothèses*

**Définition 3.7 : variante d'une formule.** Une *variante* d'une formule $A$ est une formule obtenue à partir de $A$ par changement alphabétique des variables, de telle sorte que deux occurrences de la même variable en $A$ restent des occurrences de la même variable, et que deux occurrences de variables distinctes en $A$ restent des occurrences de variables distinctes (Church, 1956, p. 86).

*Exemple.* $r \supset (s \supset r)$ et $q \supset (p \supset q)$ sont des variantes l'une de l'autre, contrairement à $p \supset (r \supset r)$ et $p \supset (p \supset p)$.

La relation "est variante de" est une relation d'équivalence[13]. Le système axiomatique étendu aux preuves à partir d'hypothèses diffère sur deux points du système initial :

1. à la règle de détachement (section 3.4.6), qui ne s'applique qu'à des thèses, on ajoute la règle du *modus ponens* qui s'applique aussi aux formes propositionnelles qui ne sont pas des thèses. Elle diffère de la règle de détachement par l'absence de *turnstile* ($\vdash$) devant les prémisses et la conclusion (cf. figure 3.14). La règle de détachement apparaît désormais

[13] Une relation binaire $\mathcal{R}$ est une relation d'équivalence si et seulement si elle remplit les trois conditions ci-dessous :

| | |
|---|---|
| $\forall x\, \mathcal{R}xx$ | Réflexivité, |
| $\forall x \forall y\, (\mathcal{R}xy \supset \mathcal{R}yx)$ | Symétrie, |
| $\forall x \forall y \forall z\, ((\mathcal{R}xy \wedge \mathcal{R}yz) \supset \mathcal{R}xz)$ | Transitivité. |

comme un cas particulier du *modus ponens*. Le détachement est une opération fermée sur les thèses (axiomes et théorèmes).

$$\begin{array}{ll} X \supset Y & \text{(prémisse majeure)} \\ X & \text{(prémisse mineure)} \\ \hline Y & \text{(conclusion)} \end{array}$$

A cette version procédurale du *modus ponens*, qui se lit : "De $X \supset Y$ et de $X$, dérivez $Y$", nous ajouterons la version déclarative suivante[14](Scott, 1974, p.148) :

$$X, X \supset Y \vdash Y$$

qui se lit : "De $X$, et de $X \supset Y$, $Y$ est dérivable". Cette version permet l'*introduction* du *turnstile* dont on aura besion pour appliquer le métathéorème de la déduction (3.4.16). Nous ajouterons également la version

$$\frac{\vdash X \supset Y}{X \vdash Y}.$$

**Figure 3.14** *Trois formes du modus ponens*

2. la règle de substitution reste en vigueur, mais elle n'est pas applicable aux prémisses qui ne sont pas des thèses.

**Définition 3.8 : la preuve à partir d'hypothèses.** Une suite finie de formules $B_1, B_2, \ldots, B_m$ est appelée *preuve à partir des hypothèses* $A_1, A_2, \ldots, A_n$ si pour chaque $i$ :

1. $B_i$ est l'une des hypothèses $A_1, A_2, \ldots, A_n$,
2. ou $B_i$ est une variante d'un axiome,
3. ou $B_i$ est inférée (par application de la règle du *modus ponens*) à partir de la prémisse majeure $B_j$ et de la prémisse mineure $B_k$ où $j < i, k < i$,
4. ou $B_i$ est inférée (par application de la règle de substitution) à partir d'une prémisse antérieure $B_j$, la variable remplacée n'apparaissant pas dans $A_1, A_2, \ldots, A_n$.

[14] D. Scott présente quatre versions du *modus ponens* qui sont toutes correctes en logique bivalente classique pour l'implication matérielle et donc compatibles (la troisième est le détachement) :

$$X, X \supset Y \vdash Y\ , \quad \frac{\vdash X \supset Y}{X \vdash Y}\ , \quad \frac{\vdash X \qquad \vdash X \supset Y}{\vdash Y}\ , \quad \frac{\vdash X}{\vdash Y}$$

Nous y ajouterons la forme du *modus ponens* de la figure 3.31, p. 150.

Une telle suite de formules — $B_m$ étant la formule finale de la suite — est appelée plus explicitement *preuve de* $B_n$ à partir des hypothèses $A_1, A_2, \ldots, A_n$, ce qu'on note par

(3.103) $A_1, A_2, \ldots, A_n \vdash B_m$

(Church, 1956, p. 87).

*Remarque.* Les variables propositionnelles figurant dans les prémisses sont en fait traitées comme des constantes.

**Définition 3.9 : la dérivation.** On appelle aussi *dérivation* ou *déduction formelle* de $B_m$ à partir de $A_1, A_2, \ldots, A_n$ ce qui a été désigné par "preuve de $B_m$ à partir des hypothèses $A_1, A_2, \ldots, A_n$".

**Définition 3.10 : la démonstration.** La démonstration est un cas particulier de dérivation; c'est une dérivation *sans* prémisse ($n = 0$).

### 3.4.12. *Exemple de preuve à partir d'hypothèses*

| | | |
|---|---|---|
| (1) | $b \vee t$ | $H1$ |
| (2) | $(b \vee a) \supset (r \wedge m)$ | $H2$ |
| (3) | $\overline{r}$ | $H3$ |
| (4) | $\overline{r} \supset (\overline{r} \vee \overline{m})$ | A. 3.35 |
| (5) | $(\overline{r} \vee \overline{m}) \equiv \neg(r \wedge m)$ | E. 3.66 |
| (6) | $((\overline{r} \vee \overline{m}) \equiv \neg(r \wedge m)) \supset ((\overline{r} \vee \overline{m}) \supset \neg(r \wedge m))$ | A. 3.38 |
| (7) | $(\overline{r} \vee \overline{m}) \supset \neg(r \wedge m)$ | M.P. 5, 6 |
| (8) | $((\overline{r} \supset (\overline{r} \vee \overline{m})) \supset (((\overline{r} \vee \overline{m}) \supset \neg(r \wedge m)) \supset (\overline{r} \supset \neg(r \wedge m)))$ | A. 3.31 |
| (9) | $((\overline{r} \vee \overline{m}) \supset \neg(r \wedge m)) \supset (\overline{r} \supset \neg(r \wedge m))$ | M.P. 8, 4 |
| (10) | $\overline{r} \supset \neg(r \wedge m)$ | M.P. 9, 7 |
| (11) | $\neg(r \wedge m)$ | M.P. 10, 3 |
| (12) | $((b \vee a) \supset (r \wedge m)) \supset \neg(r \wedge m) \supset \neg(b \vee a)$ | A. 3.41 |
| (13) | $\neg(r \wedge m) \supset \neg(b \vee a)$ | M.P. 12, 2 |
| (14) | $\neg(b \vee a)$ | M.P. 13, 11 |
| (15) | $(\neg(b \vee a) \equiv (\overline{b} \wedge \overline{a})) \supset (\neg(b \vee a) \supset (\overline{b} \wedge \overline{a}))$ | A. 3.38 |
| (16) | $\neg(b \vee a) \equiv (\overline{b} \wedge \overline{a})$ | E. 3.67 |
| (17) | $\neg(b \vee a) \supset (\overline{b} \wedge \overline{a})$ | M.P. 15, 16 |
| (18) | $\overline{b} \wedge \overline{a}$ | M.P. 17, 14 |
| (19) | $(\overline{b} \wedge \overline{a}) \supset \overline{b}$ | A. 3.32 |
| (20) | $\overline{b}$ | M.P. 19, 18 |
| (21) | $((b \vee t) \equiv (\overline{b} \supset t)) \supset ((b \vee t) \supset (\overline{b} \supset t))$ | A. 3.38 |
| (22) | $(b \vee t) \equiv (\overline{b} \supset t)$ | E. 3.75 |
| (23) | $(b \vee t) \supset (\overline{b} \supset t)$ | M.P. 21, 22 |
| (24) | $\overline{b} \supset t$ | M.P. 23, 1 |
| (25) | $t$ | M.P. 24, 20 |

### 3.4.13. *Hypothèses versus axiomes*

La raison pour laquelle la règle de substitution ne s'applique pas aux variables qui figurent dans une hypothèse (ou dans une ligne dérivée grâce à une hypothèse) est la suivante :

> Quand nous dérivons [une formule] d'une autre par un processus de substitution, la validité de notre inférence dépend de l'universalité de la prémisse [...] (Kneale et Kneale, 1961, p. 528)

Les axiomes sont des prémisses universelles. Par exemple, l'axiome

(3.104) $p \supset (p \vee q)$

est affirmé de *tout* $p$ et de *tout* $q$. Ce n'est pas le cas des hypothèses. Dans la dérivation de la section 3.4.12, la prémisse

(3.105) $b \vee t$

n'est pas affirmée de tout $b$ et de tout $t$, mais d'*au moins* un $b$ et *au moins* un $t$. En effet $b \vee t$ exprime dans la langue formelle ce qui dans la langue naturelle était exprimée par des *propositions* qui ont le statut de *constantes*, à savoir par

(3.106) Il a pris le bus ou il a pris le train.

On pourrait reformuler le contraste entre axiomes et hypothèses en disant que les premiers sont des prémisses *universelles* et les autres des prémisses *particulières*. Au lieu d'affirmer dans le *métalangage* que

(3.107) $p \supset (p \vee q)$

est une prémisse universelle et que

(3.108) $b \vee t$

est une prémisse particulière, on pourrait envisager d'écrire dans le *langage-objet*

(3.109) $\forall p \forall q\,(p \supset (p \vee q))$,

(3.110) $\exists b \exists t\,(b \vee t)$.

Il y a cependant de bonnes raisons de refuser cette manière d'exprimer le contraste entre axiomes et hypothèses :

1. On mettrait en péril l'autonomie de la logique des propositions qui étudie les dérivations reposant sur les connecteurs, à l'exclusion d'autres symboles logiques tels que les quantificateurs.
2. Si les axiomes sont traités comme des formules dont les variables propositionnelles sont liées par des quantificateurs universels, les substituts de ces variables acquièrent le statut d'*instances*, c'est-à-dire de *constantes propositionnelles*. Contrairement à la règle de substitution, la règle d'instantiation génère seulement des instances particulières de formes propositionnelles universelles, ce qui ne répond pas au but visé par les logiciens :

> Le plus souvent, les théorèmes qu'ils [les logiciens] professent d'établir en opérant des substitutions sur des variables [propositionnelles], ce sont de nouveaux énoncés universels [...] (Kneale et Kneale, 1961, p. 529).

3. Quine, enfin, fait observer que les formes propositionnelles telles que

(3.111) $\forall p(p \vee \neg p)$,

qui se lisent :

(3.112) Pour tout $p$, $p$ ou non $p$,

sont *mal formées* en cela que $p$ dans sa première occurrence (derrière "tout") appartient à la catégorie grammaticale des noms ($p$ est une variable qui admet des *noms* de phrase comme substituts et des phrases comme valeurs) et dans ses deuxième et troisième occurrences, à la catégorie grammaticale des phrases ($p$ admet des *phrases* comme substituts et des valeurs de vérité comme valeurs), ce qui prive les occurrences de $p$ dans :

(3.113) $p \vee \neg p$

de la possibilité d'être liées par le $p$ du quantificateur. Le quantificateur $\forall p$ ne lie pas plus $p$ dans la formule 3.111 que le quantificateur $\forall x$ ne le fait dans la phrase suivante :

(3.114) $\forall x(p \vee \neg p)$.

### 3.4.14. *Axiomes et schémas d'axiomes*

Puisqu'on ne peut marquer la différence de statut logique entre axiomes et hypothèses à l'aide de quantificateurs *différents* ($\forall$ pour les axiomes, $\exists$ pour les hypothèses) liant les variables propositionnelles, il faut soit imposer des *restrictions* à la *substitution*, comme le fait Church, soit faire usage d'un procédé introduit en 1927 par Von Neumann : remplacer les *axiomes* par des *schémas d'axiomes* tout en laissant les prémisses dans le langage-objet.

**Définition 3.11 : le schéma d'axiome.** On appelle schéma d'axiome un axiome dont les variables $X$, $Y$ sont des variables de formules plutôt que des variables propositionnelles. Un schéma d'axiome est, en somme, un axiome généralisé non par un quantificateur implicite, mais par le fait que la règle de substitution y est en quelque sorte *incorporée*.

Si on érige les trois axiomes du fragment implicationnel de la section 3.4.7 en *schémas d'axiome*, on se donne une infinité d'*axiomes*, à savoir toutes les

formules qui ont en commun la *structure* de ces trois schémas d'axiome :

$$\begin{array}{ll} \text{A1} & (X \supset Y) \supset ((Y \supset Z) \supset (X \supset Z)), \\ \text{A2} & ((X \supset Y) \supset X) \supset X, \\ \text{A3} & X \supset (Y \supset X). \end{array}$$

Comme on n'a plus de règle de substitution du tout, il va de soi qu'aucune substitution n'est possible sur les variables des hypothèses (sur $b$ et $t$ dans $b \vee t$ par exemple). Par ailleurs, la différence entre axiomes et hypothèses est bien marquée dans le symbolisme : le *style* des variables diffère ($X, Y, Z \ldots$[15]dans un cas ; $a, b, t, q \ldots$ dans l'autre).

Bien que les recours à la technique des *schémas d'axiome* soient fréquents dans les traités de logique, nous n'adopterons pas cette technique pour le calcul des propositions. Elle est trop abstraite et ne présente pas d'avantages pratiques qui feraient défaut à la méthode axiomatique classique. En revanche, nous utiliserons les variables $X$, $Y$ et $Z$ dans la formulation des *règles d'inférence* qui sont des fragments de *schémas de preuve*.

### 3.4.15. *Preuve à partir d'hypothèses auxiliaires*

Dans la section 3.4.10, notre objectif était de prouver la validité du raisonnement 3.83, c'est-à-dire de prouver l'énoncé métalinguistique suivant[16] :

(3.115) $(b \vee t), (b \vee a) \supset (r \wedge m), \overline{r} \vdash t$;

nous y sommes parvenus par un *détour*, à savoir en prouvant la forme implicative correspondante :

(3.116) $\vdash ((b \vee t) \wedge ((b \vee a) \supset (r \wedge m)) \wedge \overline{r}) \supset t$.

A la section 3.4.11, nous avons vu que l'on peut étendre la méthode axiomatique de manière à pouvoir démontrer non seulement que certaines formes propositionnelles sont des *théorèmes*, mais aussi que certaines formes propositionnelles sont *dérivables* de prémisses quelconques. Ceci nous a permis de démontrer *directement* 3.115 à l'intérieur du système axiomatique. La démonstration de la section 3.4.12 réclamait nettement moins d'effort inventif que la démonstration de la section 3.4.10. On est donc amené à se demander si l'on ne pourrait pas utiliser la preuve à partir d'hypothèses même là où l'objectif est de démontrer un théorème. Dans la section 3.4.10, nous voulions justifier le raisonnement 3.84 et nous avons démontré 3.116 à la place de 3.115. A présent il s'agit de faire l'inverse. Voulant démontrer le *théorème* 3.116 et non plus justifier le raisonnement 3.84, nous posons à titre de *prémisses auxiliaires* la suite de propositions

[15] On appelle métavariables les variables du métalangage.

[16] Cet énoncé se lit : "$t$" est dérivable de "$(b \vee t)$", "$(b \vee a) \supset (r \wedge m)$", "$\overline{r}$".

(3.117) $b \vee t\ ,\ (b \vee a) \supset (r \wedge m)\ ,\ \overline{r}.$

Nous essayons d'en déduire $t$ et, ayant réussi à le faire dans le cadre du système axiomatique élargi, nous absorbons les prémisses auxiliaires, qui cessent d'être des *hypothèses* pour devenir des *antécédents* de propositions hypothétiques, en application du métathéorème de la déduction qui s'énonce :

(3.118) Si $B$ est dérivable des prémisses $A_1, A_2, \ldots, A_n$, alors $A_n \supset B$ est dérivable des prémisses $A_1, A_2, \ldots, A_{n-1}$

et, formellement :

$$(3.119) \qquad \frac{A_1, A_2, \ldots, A_n \vdash B}{A_1, A_2, \ldots, A_{n-1} \vdash A_n \supset B}$$

Pour $n = 1$ on a, en particulier,

(3.120) Si $B$ est dérivable de $A$, alors $A \supset B$ est un théorème,

c'est-à-dire :

$$(3.121) \qquad \frac{A \vdash B}{\vdash A \supset B}$$

Le passage de 3.115 à 3.116 se fait au moyen de trois applications successives du métathéorème de la déduction 3.119 :

$$(3.122) \qquad \begin{gathered} b \vee t, (b \vee a) \supset (r \wedge m), \overline{r} \vdash t, \\ b \vee t, (b \vee a) \supset (r \wedge m) \vdash (\overline{r} \supset t), \\ b \vee t, \vdash ((b \vee a) \supset (r \wedge m)) \supset (\overline{r} \supset t), \\ \vdash (b \vee t) \supset (((b \vee a) \supset (r \wedge m)) \supset (\overline{r} \supset t)). \end{gathered}$$

Il suffit ensuite d'appliquer la règle d'importation deux fois successivement pour obtenir 3.116 (voir note 20, p. 153).

### 3.4.16. *Le métathéorème de la déduction*

Le métathéorème de la déduction, dû à Herbrand (1930) et Tarski, est une *méthode générale* permettant de construire une dérivation de la forme :

$$A_1, \ldots, A_{n-1} \vdash A_n \supset B,$$

à partir d'une dérivation de la forme :

$$A_1, \ldots, A_n \vdash B.$$

Au lieu de démontrer ce métathéorème (Church, 1956, pp. 88–89), nous l'illustrerons par un exemple. Supposons que nous ayons une déduction $\delta_1$ de $r$ à partir de l'ensemble d'hypothèses $\Gamma = \{p \supset (q \supset r), q, p\}$. Nous allons appliquer cette méthode pour obtenir une déduction $\delta_2$ de $p \supset r$ à partir du sous-ensemble de prémisses $\Gamma \backslash \{p\} = \{p \supset (q \supset r), q\}$.

Nous avons au départ la déduction $\delta_1$ (cf. 3.123), où $p, q, r$ sont des variables de formules (les dérivations 3.123, 3.124 et 3.125 se déroulent en effet dans le métalangage) :

(3.123)

| | | |
|---|---|---|
| (1) | $p \supset (q \supset r)$ | $H1$ |
| (2) | $q$ | $H2$ |
| (3) | $p$ | $H3$ |
| (4) | $q \supset r$ | M.P. 1,3 |
| (5) | $r$ | M.P. 4,2 |

Nous commençons "par préfixer à chacune des formules de la déduction donnée les symboles $p \supset$, en ajoutant les parenthèses nécessaires" (Kleene 1967, 1972, p. 48). Nous obtenons la déduction $\delta_2$ :

(3.124)

| | | |
|---|---|---|
| (1′) | $p \supset (p \supset (q \supset r))$ | |
| (2′) | $p \supset q$ | |
| (3′) | $p \supset p$ | théorème (cf. note 13) |
| (4′) | $p \supset (q \supset r)$ | $H1$ |
| (5′) | $p \supset r$ | |

La série de lignes de 3.124 n'est pas une déduction de (1′) à (5′) à partir des prémisses $p \supset (q \supset r)$ et $q$, mais peut être transformée en une telle déduction, si nous interpolons dans cette série des instances d'axiomes ou de théorèmes[17] judicieusement choisis.

Nous obtenons alors la liste de formules 3.125 qui, cette fois, est bien une dérivation ($\delta_3$) à partir de H1 et de H2. Nous retrouvons les formules de 3.124, que nous soulignons :

(3.125)

| | | |
|---|---|---|
| (1″) | $(p \supset (q \supset r)) \supset (p \supset (p \supset (q \supset r)))$ | A. 3 |
| (2″) | $\underline{p \supset (q \supset r)}$ | $H_1$ |
| (3″) | $\underline{p \supset (p \supset (q \supset r))}$ | M.P. 1″, 2″ |
| (4″) | $q \supset (p \supset q)$ | A. 3 |
| (5″) | $\underline{p \supset q}$ | M.P. 4″, $H_2$ |
| (6″) | $\underline{p \supset p}$ | Th. cf note 17 |
| (7″) | $(p \supset (q \supset r)) \supset (q \supset (p \supset r))$ | Th. cf note 17 |
| (8″) | $q \supset (p \supset r)$ | M.P. 7″, $H_1$ |
| (9″) | $\underline{p \supset r}$ | M.P. 8″, $H_2$ |

[17] Travaillant dans le système axiomatique de Whitehead et Russell (section 3.4.8), on démontre aisément les deux théorèmes suivants :

| | | |
|---|---|---|
| Th9. | $(p \supset (q \supset r)) \supset (q \supset (p \supset r))$ | [théorème 3.4] |
| Th10. | $p \supset p$ | [exercice] |

La suite 3.125 est bien une dérivation à partir de l'ensemble d'hypothèses $\Gamma\backslash\{p\} = \{p \supset (q \supset r), q\}$. Chaque ligne est soit un théorème, soit une ligne obtenue par la seule règle du *modus ponens* à partir d'une des deux prémisses et d'un théorème.

L'utilité pratique du métathéorème de la déduction tient en ceci :

> nous [pouvons] l'utiliser pour conclure que certaines démonstrations *existent* sans nous imposer de les construire effectivement (Kleene, 1967-1972, p. 50).

## 3.5. La méthode de la déduction naturelle

En 1934, Gentzen et Jaskowski ont, indépendamment l'un de l'autre, inventé une méthode de déduction connue aujourd'hui sous le nom de déduction naturelle dans laquelle il n'y a pas d'axiome, mais uniquement des règles d'inférence. Cette méthode présente un triple avantage par rapport à la méthode axiomatique :

1. Elle est d'un maniement nettement plus aisé que la méthode axiomatique et imite, dans une certaine mesure, la manière de raisonner spontanée du mathématicien.
2. Elle permet de rendre compte de la propriété qu'ont certaines preuves déductives de propager l'évidence.
3. Elle réussit à expliquer le sens des connecteurs pris isolément.

Ce sont autant de raisons qui justifient ici une étude détaillée de cette méthode et une présentation qui fournisse au lecteur toutes les informations dont il a besoin pour la mettre en pratique.

### 3.5.1. *Fragment implicationnel en déduction naturelle*

**Définition 3.12 : la dérivation dans le fragment implicationnel.** On appelle dérivation dans le fragment implicationnel de la déduction naturelle une suite finie de lignes auxquelles sont associés :

1. un numéro d'ordre,
2. un ensemble — éventuellement vide — de *numéros de prémisse* désignant les prémisses dont la ligne *dépend*, les numéros de prémisse sont entourés d'accolades,
3. un nom de règle d'inférence et les numéros des lignes servant d'entrée à l'inférence.

Cette suite finie est construite par application des trois règles ci-dessous (adaptée de Mates, 1965, p.93) :

*Règles*

**P (Introduction de prémisses).** N'importe quelle forme propositionnelle peut être introduite à titre de prémisse. Une prémisse introduite par cette règle ne dépend que d'elle-même. Elle prendra donc comme *numéro de prémisse* son numéro d'ordre, c'est-à-dire $n$ si elle est la $n$-ième ligne.

**M.P. (Modus ponens).** $Y$ peut être introduit comme ligne dans une dérivation à condition que $X$ et $X \supset Y$ y figurent déjà. $Y$ hérite de toutes les prémisses de $X$ et de $X \supset Y$ et dès lors prend comme numéros de prémisses tous les numéros de prémisses de $X$ et de $X \supset Y$.

**C (Introduction d'un conditionnel).** Le conditionnel $X \supset Y$ peut être introduit si $Y$ apparaît dans une ligne antérieure. $X \supset Y$ hérite de toutes les prémisses et dès lors de tous les numéros de prémisse de $Y$, à l'exception du numéro de la ligne (si elle existe) sur laquelle $X$ apparaît. Soustraire de l'ensemble des numéros de prémisse de $X \supset Y$ celui de la ligne où $X$ apparaît, cela s'appelle *décharger* la prémisse $X$.

1. Décharger une prémisse est une opération permise, mais non obligatoire.
2. On décharge les prémisses dans l'ordre inverse de celui dans lequel on les a introduites.
3. Une fois qu'une prémisse est déchargée, on ne peut plus rien déduire de cette prémisse, ni d'une ligne qui aurait nécessité cette prémisse pour sa déduction.

**Définition 3.13 : la démonstration en déduction naturelle.** On appelle démonstration une dérivation à partir de l'ensemble vide de prémisses que l'on note $\Lambda$.

**Définition 3.14 : le théorème en déduction naturelle.** $X$ est un théorème si et seulement si $X$ est dérivable de $\Lambda$.

### 3.5.2. *Exemple de démonstration en déduction naturelle*

Redémontrons à nouveau le théorème 3.1 de la section 3.4.7

*Théorème* 3.1 $((q \supset p) \supset r) \supset (p \supset r)$

*Remarque.* La démonstration du théorème 3.1 (cf section 3.4), faite à l'aide du métathéorème de la déduction, mêlait les formules du langage-objet (1-5) et des propositions du métalangage (6-8). La démonstration de ce théorème en déduction naturelle, au contraire, *se déroule tout entière dans le langage-objet*. Elle est donc plus directe et plus naturelle. On ne discourt pas sur la dérivabilité de certaines formules à partir d'autres, on *effectue* la dérivation.

| | | | |
|---|---|---|---|
| 1 | {1} | $(q \supset p) \supset r$ | $P^1$ |
| 2 | {2} | $p$ | $P^2$ |
| 3 | {2} | $q \supset p$ | C(2) [sans décharge] |
| 4 | {1,2} | $r$ | M.P(1,3) |
| 5 | {1} | $p \supset r$ | C(2,4) [décharge de $P^2$] |
| 6 | $\Lambda$ | $((q \supset p) \supset r) \supset (p \supset r)$ | C(1,5) [décharge de $P^1$] |

**Figure 3.15** *Démonstration du théorème 3.1 en déduction naturelle*

Quand on applique la règle d'*introduction de prémisse*, on *suppose* admise une proposition. Quand on *décharge* une hypothèse $p$ grâce à l'introduction d'un conditionnel, on remplace la forme propositionnelle $p$ qui était affirmée à titre de supposition, par une forme propositionnelle de *forme* conditionnelle, mais *affirmée* inconditionnellement.

Ces distinctions ont-elles un sens pour une langue formelle non interprétée ? Kirwan le conteste :

> [...] quand on étudie la déduction naturelle, on est incité à "supposer [*assume* en anglais]" n'importe quelle formule, disons "$p \wedge q$" à titre de "prémisse" pour en déduire la "conclusion" "$p$". Mais cela est non satisfaisant. [...] Si $p$ est une formule non interprétée, on ne peut pas littéralement la supposer ou la déduire. Vous pourriez tout autant supposer une tache d'encre (Kirwan, éd. par Smiley, 1982-83, p.13)

L'objection de Kirwan n'est pas dirimante. Il y a moyen de distinguer les *formules supposées* des *formules assertées* en recourant à des critères purement formels (le fait d'être précédé ou non d'un ensemble non vide de numéros de prémisses).

Sur le plan heuristique, la démonstration présentée à la figure 3.15 ne fait appel à aucun axiome. *Toutes* les parties de la démonstration sont des constituants de la formule à démontrer et la structure de cette dernière donne plusieurs indications sur la manière de construire la démonstration. Ces bonnes propriétés ne sont cependant pas conservées quand on étend la méthode de la déduction naturelle à tout le calcul des propositions.

*Exercice.*

Démontrer :

(3.126) $(p \supset (q \supset r)) \supset ((r \supset s) \supset (p \supset (q \supset s))$.

### 3.5.3. *Extension de la déduction naturelle*

Pour représenter les dérivations, nous adopterons un formalisme plus courant et plus maniable que celui de Mates, mais, hélas, aussi plus *artificiel*, en ce sens qu'il utilise la notion *procédurale* de "dérivation dans la *portée* d'une prémisse auxiliaire $P_i$" au lieu de la notion *logique* de "dérivation sous la *dépendance logique* d'une prémisse auxiliaire $P_i$". Cet *artifice* n'affecte pas la puissance déductive de la méthode et il permet d'écrire des dérivations plus lisibles.

**Définition 3.15 : la dérivation.** On appelle *dérivation* en déduction naturelle une suite finie de lignes chacune précédée d'un numéro d'ordre et suivie du nom de la règle qui justifie la ligne considérée, ainsi que des numéros de lignes servant d'entrée (*input*) à la règle. La portée des prémisses auxiliaires ($P^i$) sera indiquée par la zone délimitée par une flèche ou deux flèches comme dans la figure 3.18 (adaptée de Johnstone et Anderson 1961).

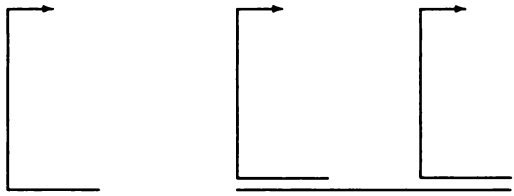


**Figure 3.16** *Marqueurs de portée des prémisses auxiliaires*

*Remarque.* Dans une dérivation, convenons d'associer aux *prémisses ordinaires* des chiffres romains et d'associer les symboles $P^1, \ldots, P^n$ aux *prémisses auxiliaires*, c'est-à-dire aux prémisses introduites par la règle I($\supset$) et destinées à être déchargées par la règle E($\supset$) (voir figure 3.17).

### 3.5.4. *Exemple de dérivation et de démonstration*

Toute dérivation peut être transformée en démonstration. Il suffit de traiter les *prémisses* comme *prémisses auxiliaires*, c'est-à-dire de les décharger (cf. figure 3.18).

Au lieu de représenter directement le raisonnement par une dérivation en déduction naturelle (figure 3.19), on peut construire une démonstration de la forme implicative qui correspond (au sens de 1.5.4) à ce raisonnement.

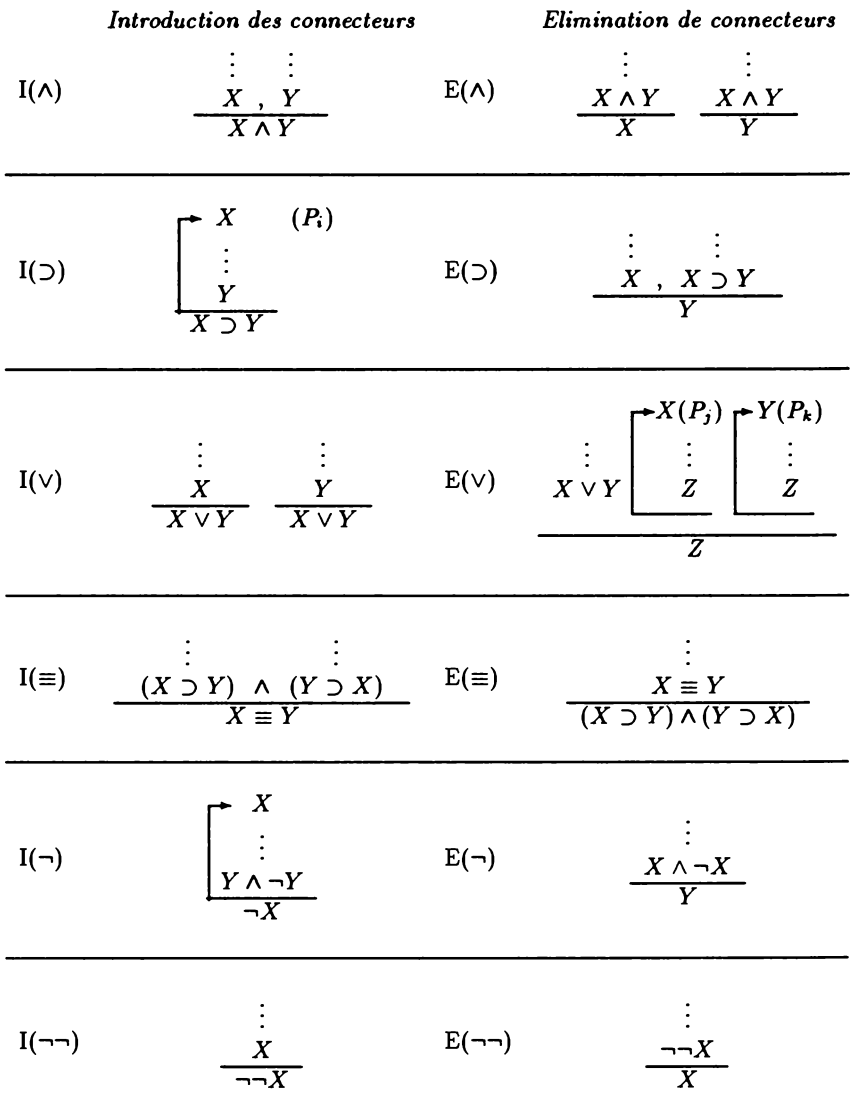



**Figure 3.17** *Règles de déduction naturelle*

### 3.5.5. *Règles facultatives en déduction naturelle*

1. *Règle d'échange*

   Si $X'$ est obtenu à partir de $X$ en remplaçant en $X$ une occurrence de

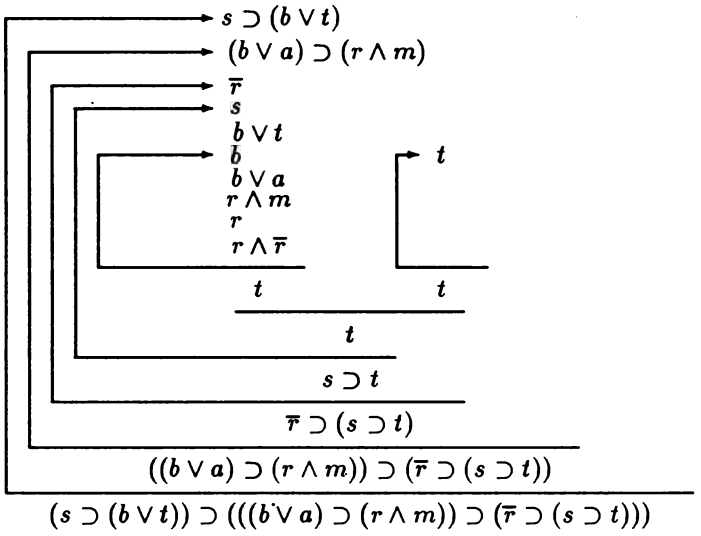



**Figure 3.18** *Démonstration de la forme implicative de 3.18 en déduction naturelle*

la formule $Y$ par une occurrence de la formule $Y'$ qui lui est équivalente en vertu des équivalences 3.65 à 3.82, alors si $X$ est une ligne dans une dérivation, $X'$ peut être introduite dans la dérivation dans la portée des mêmes prémisses auxiliaires que $X$.

2. *Introduction de théorèmes dans une dérivation*

Tout théorème déjà démontré peut être introduit comme ligne supplémentaire dans une dérivation. Comme, par définition, un théorème découle de l'ensemble vide de prémisses, l'introduction d'un théorème n'accroît pas le nombre de prémisses et le théorème ne doit donc pas être déchargé.

Le recours à des théorèmes lors de la construction d'une dérivation en déduction naturelle n'efface pas la différence entre dérivation axiomatique et dérivation naturelle. En effet, les théorèmes introduits dans une dérivation en déduction naturelle sont, en fait, des dérivations tronquées dont on n'a conservé que la dernière ligne, celle-ci ne dépendant que de l'ensemble vide de prémisses. Par exemple, le théorème $p \supset (q \supset (p \wedge q))$ inséré dans une dérivation en déduction naturelle est un raccourci de la *sous-dérivation* ou *dérivation locale* de la figure 3.20

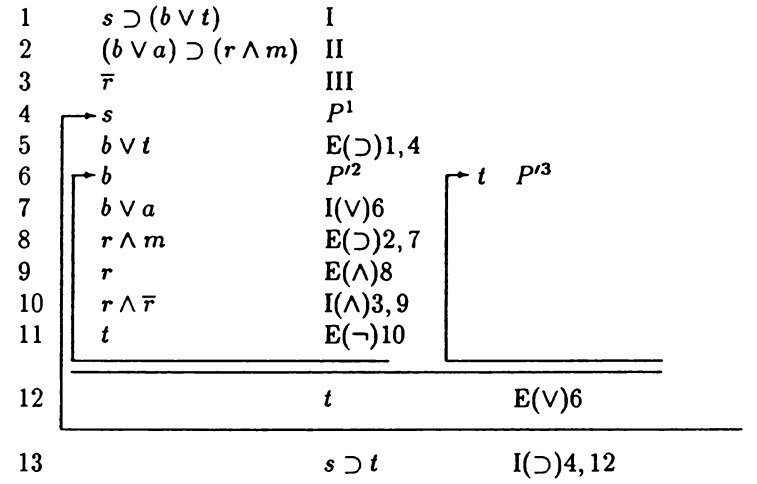



**Figure 3.19** *Dérivation de 3.18 en déduction naturelle*

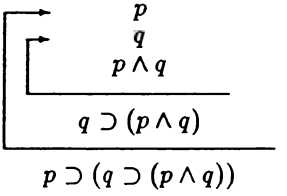



**Figure 3.20** *Théorème vu comme sous-dérivation*

3. *Cinq règles dérivées d'emploi fréquent*

(a) *Modus tollens* (M.T.) $\dfrac{X \supset Y, \quad \neg Y}{\neg X}$

(b) *Syllogisme hypothétique* (S.H.) $\dfrac{X \supset Y, \quad Y \supset Z}{X \supset Z}$

(c) *Syllogisme disjonctif* (S.D.) $\dfrac{X \vee Y, \quad \neg X}{Y}$

(d) *Dilemme constructif* (D.C.) $\dfrac{X \vee Y, \quad X \supset Z, \quad Y \supset U}{Z \vee U}$

(e) *Dilemme destructif* (D.D.) $\dfrac{\neg X \vee \neg Y, \quad Z \supset X, \quad U \supset Y}{\neg Z \vee \neg U}$

Nous nous bornerons à démontrer la cinquième règle dérivée.

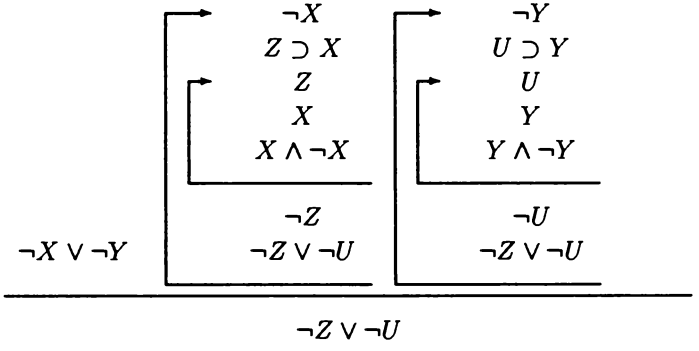



**Figure 3.21** *Dérivation du dilemme destructif*

Les règles dérivées permettent de raisonner avec beaucoup plus de souplesse. Le problème de diagnostic médical examiné ci-dessous serait sensiblement plus difficile si l'on s'imposait de renoncer aux règles dérivées.

### 3.5.6. *Application des règles facultatives*

*Lois d'un manuel de pathologie*

(3.127) Si le patient a la rougeole, il a de la fièvre (I)

(3.128) Si le patient a une hépatite, mais pas la rougeole, il a le teint jaune (II)

(3.129) Si le patient a de la fièvre ou le teint jaune, il a une hépatite ou la rougeole (III)

*Faits observés lors de l'examen clinique*

(3.130) Le patient n'a pas le teint jaune (IV)

(3.131) Le patient a de la fièvre (V)

*Problème : établir le diagnostic*

On représente l'énoncé "le patient a la rougeole" par la variable propositionnelle $r$, l'énoncé "le patient a de la fièvre" par $f$ et ainsi de suite.

(3.132)

| | | |
|---|---|---|
| 1 | $r \supset f$ | I |
| 2 | $(h \wedge \neg r) \supset j$ | II |
| 3 | $(f \vee j) \supset (h \vee r)$ | III |
| 4 | $\overline{j}$ | IV |
| 5 | $f$ | V |
| 6 | $\neg(h \wedge \neg r)$ | M.T. 2,4 |
| 7 | $\neg h \vee \neg\neg r$ | D.M. 6 |
| 8 | $\neg h \vee r$ | D.N.(7) + éch |
| 9 | $h \supset r$ | IM 8 |
| 10 | $f \vee j$ | I(∨)5 |
| 11 | $h \vee r$ | E(⊃)3, 10 |
| 12 | $r \vee h$ | COM11 |
| 13 | $\neg\neg r \vee h$ | D.N. 12 + éch |
| 14 | $\neg r \supset h$ | IM13 |
| 15 | $(\neg r \supset h) \wedge (h \supset r)$ | I(∧)9, 14 |
| 16 | $\neg r \supset r$ | S.H.15 |
| 17 | $\neg\neg r \vee r$ | IM16 |
| 18 | $r \vee r$ | D.N.17 + éch |
| 19 | $r$ | IDM18 |

### 3.5.7. *Méthode axiomatique et déduction naturelle*

Dans la méthode axiomatique classique, les deux difficultés principales sont

1. le choix d'une substitution utile parmi toutes les substitutions possibles. Ces substitutions sont en nombre infini.
2. le repérage des axiomes ou théorèmes déjà démontrés qui doivent fournir les *entrées (inputs)* à la règle de détachement. La lourdeur de cette tâche croît au fur et à mesure que la liste des théorèmes s'allonge.

Dans la méthode de déduction naturelle, on ne fait pas appel à la règle de substitution, si ce n'est dans la mesure où celle-ci est incorporée aux variables $X$, $Y$, $Z$ qui sont des variables de formules. Rappelons que les règles de la déduction naturelle — tout comme la règle du détachement dans la méthode axiomatique — ne sont pas des fragments de dérivation dans le *langage-objet*, mais des *schémas* de tels fragments dans le métalangage.

Comme on n'a plus d'axiomes et que le recours aux théorèmes déjà démontrés est facultatif, on n'est plus confronté à un choix difficile d'hypothèses. On démarre *in medias res*, à partir des prémisses données et, éventuellement, de l'antécédent de la formule à démontrer, si celle-ci est une conditionnelle.

Il faut cependant faire un effort d'imagination pour construire les démonstrations (ou les dérivations), mais l'effort est généralement moins grand que celui exigé dans la méthode axiomatique.

Bien qu'une démonstration en déduction naturelle limitée aux règles d'introduction et d'élimination des connecteurs fasse intervenir moins de *formules exogènes*, c'est-à-dire de formules qui ne sont pas des parties de la formule à dériver ou à démontrer, elle en fait cependant intervenir par le biais de la règle d'introduction du connecteur $\vee$.

$$\frac{X}{X \vee Y}$$

*Remarque.* Prawitz (1971) a démontré cependant qu'il est possible de réécrire systématiquement les démonstrations en déduction naturelle qui font usage de formules exogènes en dérivations qui opèrent uniquement sur des sous-formules du théorème à démontrer.

L'équivalence de la méthode de déduction naturelle et de la méthode axiomatique a été démontrée pour certains systèmes (cf Moens, 1980, p.49-55).

### 3.5.8. *Définition récursive de la notion de dérivation*

Désignons par $\pi$ les dérivations, par $\Delta$ les prémisses et par $\phi$ la conclusion. Chaque application d'une règle d'inférence permet de former une nouvelle dérivation à partir d'une dérivation déjà construite. On peut dès lors donner une définition inductive du prédicat à trois places $\mathcal{D}(\pi, \phi, \Delta)$ qui se lit "$\pi$ est une dérivation de la conclusion $\phi$ à partir des prémisses $\Delta$". Voici la définition inductive que donne N. Tennant dans *Natural Deduction* (Tennant, 1978, p. 50 et suivantes).

Le cas de base est la dérivation à une ligne. Dans une dérivation réduite à une ligne, $\pi$ (la dérivation) et $\phi$ (la formule déduite) se confondent et comme la formule déduite n'a d'autre prémisse qu'elle-même, l'ensemble $\Delta$ est le singleton $\{\phi\}$. Formellement, cela s'écrit :

1. *Base :*

$$\mathcal{D}(\phi, \phi, \{\phi\})$$

2. *Clause inductive :* à partir de $n$ dérivations, on contruit une $(n+1)^{ième}$ dérivation $D$ qui intègre les précédentes. Une dérivation globale[18] peut comporter $n$ branches parallèles. C'est le cas si elle contient des dilemmes emboîtés (voir l'application de 3 dans 5.4.8). Faire un pas dans une dérivation, c'est franchir la barre horizontale d'inférence et ajouter une

[18] L'opposition entre "dérivation locale" et "dérivation globale" est relative. Une dérivation globale peut être locale par rapport à une dérivation dans laquelle elle est enchâssée. Réciproquement, une dérivation locale peut être globale par rapport à une dérivation qu'elle contient.

formule sous cette barre, qui est le point d'aboutissement des dérivations locales $\pi_1$ à $\pi_n$. Lors du franchissement de la barre d'inférence, l'ensemble des prémisses est généralement modifié, soit qu'il s'accroisse par le regroupement des ensembles de prémisses $\Delta_1, \ldots, \Delta_n$, soit qu'il diminue à la faveur d'une décharge d'hypothèse auxiliaire. En tout état de cause, le nouvel ensemble de prémisses sera une fonction $f$ des dérivations antérieures $\pi_1, \ldots, \pi_n$ et de la formule $\phi$. Formellement, cela s'écrit :

$$\text{si } \left\{ \begin{array}{c} \mathcal{D}(\pi_1, \phi_1, \Delta_1) \\ \vdots \\ \mathcal{D}(\pi_n, \phi_n, \Delta_n) \end{array} \right\} \text{ alors } \mathcal{D}\left( \frac{\pi_1 \ldots \pi_n}{\phi}, \phi, f(\pi_1 \ldots \pi_n, \phi) \right)$$

où $f(\pi_1 \ldots \pi_n, \phi_n)$ désigne l'ensemble effectivement déterminable de prémisses dont dépend la nouvelle conclusion $\phi$. Dans beaucoup de cas, ce sera la réunion des ensembles de prémisses $(\Delta_1 \cup \ldots \cup \Delta_2)$. Parfois ce sera le reste après soustraction des prémisses déchargées.

3. *Clause de fermeture :* rien n'est une dérivation, sauf si elle découle de la clause de base ou de clauses qui sont des instances de la clause inductive.
4. *Définition de la preuve :* on appelle preuve ou démonstration toute dérivation *sans* prémisse. En d'autres termes, l'ensemble $\Delta$ des prémisses est vide et toutes les prémisses auxiliaires ont été déchargées.

*Remarques.*

1. Les lignes de toute dérivation sont en nombre fini.
2. Chaque règle d'inférence spécifie les prémisses dont elle dépend.
3. La réponse à la question de savoir quelle règle a été appliquée et si elle l'a été correctement est déterminée par des considérations de structure seulement.
4. Quand on prolonge une dérivation par l'adjonction d'une inférence, le stock des prémisses est accru : $\Delta_1 \cup \Delta_2 \ldots$, sauf quand on décharge une hypothèse auxiliaire, auquel cas il est diminué : $\Delta \setminus \{\psi\}$.
5. La dérivation *globale* est faite de dérivations *locales*. L'enchâssement des dérivations locales dans la dérivation globale n'est pas quelconque. Chaque franchissement de la barre d'inférence doit être autorisé par l'une des règles d'inférence citées. Les erreurs des novices sont souvent dues à la méconnaissance de cette exigence. Une erreur fréquente est la simplification dans une formule qui n'est pas autonome.

   Par exemple,

$$\frac{(1) \quad r \supset (p \wedge q)}{(2) \quad r \supset p}$$

Bien que la ligne 2 découle logiquement de la ligne 1, elle n'en découle pas *en vertu* de la seule règle d'élimination du $\wedge$. Voici un exemple plus grave de simplification illicite, où la formule dérivée n'est pas une conséquence logique de la formule initiale :

$$\frac{\neg(p \wedge q)}{\neg p}$$

**Figure 3.22** *Simplification fallacieuse*

## 3.6. Evidence, sens et notation

### 3.6.1. *L'évidence des preuves en déduction naturelle*

Les fragments de preuve constitués par l'application de *règles d'introduction* de connecteurs possèdent une *évidence intrinsèque*. Celles-ci formulent, en effet, les *conditions suffisantes* de l'assertion correcte d'une formule moléculaire en fonction des conditions d'assertion correcte des atomes. Par exemple, la règle d'introduction du $\wedge$ nous dit que pour être en droit d'affirmer $X \wedge Y$, il suffit d'être en droit d'affirmer $X$ et d'être en droit d'affirmer $Y$. En d'autres termes, disposer d'un argument ou d'une preuve de $X$ et d'un argument ou d'une preuve de $Y$ est une condition suffisante pour avoir un argument ou une preuve de $X \wedge Y$.

Alors que les règles d'introduction de propositions d'une certaine forme énoncent les conditions permettant d'inférer une proposition de la forme en question, les règles d'élimination d'une certaine forme énoncent les conclusions immédiates qui peuvent être tirées de propositions de la forme en question.

Mettons, par exemple bout à bout la règle d'introduction de $\supset$ et la règle d'élimination de $\supset$ :

$$\dfrac{\dfrac{\begin{array}{c}[X]\\ \vdots \\ Y\end{array}}{X \supset Y} \quad X}{Y}$$

**Figure 3.23** *Introduction suivie d'élimination*

Pour prouver l'énoncé hypothétique $X \supset Y$, on a construit une preuve hypothétique de $Y$ à partir de $X$ supposé vrai, à titre de prémisse auxiliaire.

Mettons-nous à présent dans la situation où nous avons une preuve de la mineure $X$, et réactivons la preuve de $Y$ à partir de $X$. Nous pouvons conclure $Y$ directement, c'est-à-dire sans passer par l'énoncé conditionnel $X \supset Y$ qui apparaît comme un *détour* inutile (figure 3.24).

$$\dfrac{\dfrac{\vdots}{[X]}}{\dfrac{\vdots}{Y}}$$

**Figure 3.24** *Elimination des détours pour le conditionnel*

Cette démarche peut être accomplie pour tous les connecteurs, comme l'affirme le *principe d'inversion* implicite chez Gentzen et formulé pour la première fois par Lorenzen en 1950. Nous citerons ici la formulation de Prawitz :

(3.133) Soit $\alpha$ une application d'une règle d'élimination qui a $B$ comme conséquence. Alors, les déductions qui satisfont la condition suffisante [...] pour dériver la prémisse majeure de $\alpha$ [celle qui contient le connecteur *introduit*], "contiennent" déjà une déduction de $B$, quand elles sont combinées avec des déductions des prémisses mineures de $\alpha$ si tant est qu'il y en ait. La déduction de $B$ est alors susceptible d'être obtenue directement à partir des déductions données sans l'addition de $\alpha$. (Prawitz, 1965, p. 33)

Nous avons décrit l'opération justificative pour le $\supset$ dans la section précédente. Il en existe aussi pour le $\wedge$ et le $\vee$. L'opération justificative pour le $\wedge$ permet de *transformer* une introduction de ce connecteur suivie de son élimination (1) en une déduction *sans détour* (2) (voir figure 3.25).

$$(1)\ \dfrac{\dfrac{\vdots}{X_1} \quad \dfrac{\vdots}{X_2}}{\dfrac{X_1 \wedge X_2}{X_i}} \qquad (2)\ \dfrac{\vdots}{X_i}$$

**Figure 3.25** *Elimination des détours pour la conjonction*

L'opération justificative pour le $\vee$ permet de transformer (1) en (2) (voir figure 3.26).

On qualifie de *normale* une preuve en déduction naturelle dans laquelle toutes les occurrences de propositions figurant simultanément comme conclusion d'une introduction et comme prémisse majeure d'une élimination ont disparu. Prawitz a démontré le *théorème de la forme normale* suivant (Prawitz,

1965) : "Si $\Gamma \vdash A$, alors il existe une déduction dans laquelle aucune occurrence de formule n'est la conséquence d'une application d'une règle d'introduction et en même temps la prémisse majeure d'une règle d'élimination."

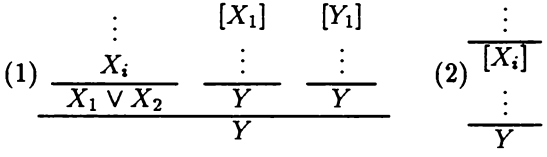


**Figure 3.26** *Elimination des détours pour la disjonction*

### 3.6.2. *Signification des constantes logiques*

On peut résumer dans la table ci-dessous le sens de $\vee$, $\wedge$ et $\supset$.

| | | |
|---|---|---|
| L'énoncé "$A \vee B$" est vrai | pourvu que | "$A$" soit vrai ou "$B$" soit vrai |
| L'énoncé "$A \wedge B$" est vrai | pourvu que | "$A$" soit vrai et "$B$" soit vrai |
| L'énoncé "$A \supset B$" est vrai | pourvu que | "$B$" soit vrai pourvu que "$A$" soit vrai |

**Figure 3.27** *Définition sémantique des constantes logiques*

Martin-Löf suggère de faire faire un quart de tour à cette table dans le sens opposé à celui des aiguilles d'une montre. On obtient alors les règles d'inférence de le figure 3.28 :

$$\frac{\text{"}A\text{" est vrai}}{\text{"}A \vee B\text{" est vrai}} \qquad \frac{\text{"}B\text{" est vrai}}{\text{"}A \vee B\text{" est vrai}}$$

$$\frac{\text{"}A\text{" est vrai} \quad \text{"}B\text{" est vrai}}{\text{"}A \wedge B\text{" est vrai}} \qquad \frac{\begin{array}{c}[\text{"}A\text{" est vrai}] \\ \text{"}B\text{" est vrai}\end{array}}{\text{"}A \supset B\text{" est vrai}}$$

**Figure 3.28** *Définition inférentielle des constantes logiques*

Ce changement de perspective révèle que

> Si vous interprétez les conditions de vérité de cette manière, vous voyez qu'elles sont identiques aux règles d'introduction pour les constantes logiques telles qu'elles ont été formulées par Gentzen (Martin-Löf, *Ibid.*, p. 411).

Gentzen a suggéré que les règles d'introduction des constantes logiques pourraient bien être considérées comme des ***définitions*** des constantes logiques :

> A chaque symbole logique $\wedge, \vee, \ldots, \supset, \neg$, appartient précisément une figure d'inférence qui 'introduit' un symbole — comme symbole terminal d'une formule — et une figure qui l''élimine'. Le fait que les figures d'inférence $\wedge$-Elim et $\vee$-Intr ont chacune deux formes constitue une dérivation triviale, purement externe et sans intérêt. Les introductions représentent, en quelque sorte, les 'définitions' des symboles en cause, et les éliminations ne sont pas plus, en dernière analyse, que les conséquences de ces définitions. Ce fait peut être exprimé comme suit : en éliminant un symbole, nous pouvons utiliser la formule avec le symbole terminal en cause seulement 'dans le sens conféré à ce dernier par la règle d'introduction de ce symbole'. Un exemple peut éclairer ce que cela veut dire : nous fûmes en mesure d'introduire la formule $\mathcal{A} \supset \mathcal{B}$ là où il existait une dérivation de $\mathcal{B}$ à partir de la formule supposée de $\mathcal{A}$. Si nous désirions utiliser cette formule en éliminant le symbole $\supset$ [...] nous pourrions le faire précisément en inférant $\mathcal{B}$ directement, une fois que $\mathcal{A}$ a été prouvé, car ce qu'atteste $\mathcal{A} \supset \mathcal{B}$, c'est précisément l'existence d'une dérivation de $\mathcal{B}$ à partir de $\mathcal{A}$.
>
> En rendant ces idées plus précises il devrait être possible de présenter les inférences E [règles d'élimination] comme des fonctions uniques des inférences I [règles d'introduction] correspondantes, sur la base de certaines exigences (Gentzen, 1934, 1968, p. 80–81).

Plusieurs auteurs ont entrepris de mener à son terme le programme esquissé par Gentzen dans le passage précité : définir les constantes logiques par les règles de déduction naturelle.

Pour mener à bien ce projet, il ne suffisait pas de *compléter* ce qui était ébauché dans le texte fondateur de Gentzen, il fallait, par surcroît, *étendre* le système de la déduction naturelle. C'est ce qu'a fait Peter Schröder-Heister dans *A Natural Extension of Natural Deduction* (1984).

L'innovation majeure de P. Schröder-Heister, c'est d'avoir montré que l'on peut introduire localement dans une dérivation, non seulement des formules *ad hoc* à titre de prémisses auxiliaires que l'on décharge ensuite, mais que l'on peut faire cela aussi avec des règles de niveau supérieur. En d'autres termes, on peut avoir comme prémisses auxiliaires dans une règle de niveau $n$ une autre règle de niveau $n+1$ qui pourra être déchargée exactement comme une prémisse auxiliaire. L'innovation de Schröder-Heister nous donne le moyen d'*apparier* les règles d'*introduction* aux règles d'*élimination* comme le suggère Gentzen et de "dériver" (Sundholm, 1986, p. 496) les secondes des premières. Ceci exige que l'on reformule autrement les règles d'élimination de $\wedge$, $\vee$ et $\supset$.

*L'enjeu philosophique*

Que l'on définisse les constantes logiques par l'intermédiaire des axiomes dans lesquels elles figurent comme termes primitifs, ou par les biais de règles de preuve, la démarche s'accomplit de toute manière dans le cadre de la théorie de la preuve et non dans le cadre de la sémantique. Il y a un enjeu philosophique associé à ce choix : on est amené à se poser la question suivante : est-ce la notion de preuve ou la notion de vérité qui est conceptuellement première ? (Sur cette question, voir Engel, 1989.)

Les logiciens intuitionnistes font observer que la *prétention* à la vérité ne peut être séparée de la prétention à la *connaissance*, ce qui conduit à un concept de vérité qui dépend de la connaissance. Selon l'analyse intuitionniste de la notion de vérité,

> [...] c'est le concept de savoir qu'une proposition est vraie, c'est-à-dire le concept de connaître une preuve de la proposition qui est la notion conceptuellement première, et alors on extrait de cette notion celle de vérité en disant qu'une proposition est vraie si elle est directement prouvable, c'est-à-dire si elle peut être prouvée par les moyens les plus directs. En outre, les conditions de vérité des constantes logiques qui ont la même formulation que celle à laquelle nous sommes accoutumés, sont interprétées [nous l'avons vu au début de cette section] conformément à la manière dont elles apparaissent comme conditions de preuve directe (Martin-Löf, *Op. Cit.*, p. 414).

Il ne faut pas se méprendre sur l'affirmation de Martin-Löf. Celui-ci ne nie pas l'existence de vérités indémontrables dans un système – il ne met pas en cause le théorème de Gödel établissant qu'il y a des vérités de l'arithmétique indémontrables à partir des axiomes de Peano. Ces propositions non dérivables des axiomes de Peano ou de tout autre ensemble récursif d'axiomes figurent parmi les vérités de l'arithmétique non formalisée. Ce que Martin-Löf met en doute, c'est l'existence de vérités qui seraient *en principe* inconnaissables[19].

### 3.6.3. *Variante notationnelle utile de la déduction naturelle*

Prenons, au hasard, l'une des règles d'introduction ou d'élimination des connecteurs, par exemple, la règle d'élimination du connecteur $\supset$ (figure 3.29)

[19] L'*axiome du choix* serait peut-être un exemple. Cet axiome affirme que dans chacun des sous-ensembles disjoints non vides $M, N, R, \ldots$ d'un ensemble $t$, il est toujours possible de choisir un membre unique $m, n, r, \ldots$ et de rassembler ces derniers en un ensemble (Fraenkel et al. 1973, p. 55).

Paul Cohen a démontré en 1963 que l'axiome du choix est *indépendant* des axiomes de la théorie des ensembles de Zermelo-Fraenkel. Des résultats décisifs mais incomplets concernant ce problème sont dus à Fraenkel (1922), Gödel (1938), Specker, Mendelson et Schoenfield (1951–1955).

qui précisait que si $X$ et $X \supset Y$ étaient des *thèses*, $Y$ en était une aussi.

$$\frac{X \qquad X \supset Y}{Y}$$

**Figure 3.29** *Première version de la règle d'élimination de* $\supset$

Ecrit de la sorte, le schéma d'inférence est incomplet. Il ne nous dit rien sur le statut logique des formules qui prendront la place de $X$, $X \supset Y$ et $Y$, contrairement à la règle de détachement du calcul des propositions décrite à la figure 3.30.

$$\frac{\vdash X \qquad \vdash (X \supset Y)}{\vdash Y}$$

**Figure 3.30** *Deuxième version de la règle d'élimination de* $\supset$

Une formulation tout à fait explicite de la règle d'élimination du connecteur $\supset$ en déduction naturelle serait celle de la figure 3.31, qui se lit : "de $X$ dérivable

$$\frac{\Gamma \vdash X \qquad \Delta \vdash (X \supset Y)}{\Gamma, \Delta \vdash Y}$$

**Figure 3.31** *Troisième version de la règle d'élimination de* $\supset$

de l'ensemble de formules $\Gamma$ et de $X \supset Y$ dérivable de l'ensemble des formules $\Delta$, déduisez que $Y$ est dérivable des ensembles de formules $\Gamma, \Delta$". Il n'est donc plus nécessaire que les prémisses de la règle soient des *thèses* (théorèmes ou axiomes) comme c'était le cas dans la règle de détachement formulée plus haut. Il n'est pas même nécessaire que les prémisses de la règle soient des *énoncés affirmés*. Il suffit que $X$ et $X \supset Y$ soient dérivables, *fût-ce à partir de suppositions*. Compléter la formulation initiale comme nous venons de la faire est une nécessité si nous voulons pouvoir faire place aux preuves conditionnelles qui ne partent ni d'axiomes, ou de théorèmes, ni même de simples affirmations, mais de propositions qui sont "supposées" plutôt qu'affirmées.

Les expressions de la forme $\vdash (X \supset Y)$ seront réécrites $X \Rightarrow Y$ et appelées "séquents". La portée de ce changement de notation sera précisée dans le chapitre 4.

Reformulons les règles de la déduction naturelle à l'aide de ce nouveau symbole (Thayse et al. 1988). Nous dirons ensuite quel est *l'intérêt pratique* de cette opération qui, à première vue, paraît n'être qu'une excroissance notationnelle inutile.

L'intérêt de reformuler la déduction naturelle à l'aide d'un système où le terme primitif est le prédicat $\Rightarrow$ tient en ceci : cette nouvelle présentation

*Règles de base*

1 Réflexivité $\dfrac{}{X \Rightarrow X}$

2 Atténuation $\dfrac{\Gamma \Rightarrow X}{\Gamma, \Delta \Rightarrow X}$

*Règles d'introduction*

3 $(\wedge)$ $\dfrac{\Gamma \Rightarrow X \quad \Gamma \Rightarrow Y}{\Gamma \Rightarrow X \wedge Y}$

4 $(\vee)$ $\dfrac{\Gamma \Rightarrow X}{\Gamma \Rightarrow X \vee Y} \qquad \dfrac{\Gamma \Rightarrow Y}{\Gamma \Rightarrow X \vee Y}$

5 $(\supset)$ $\dfrac{\Gamma, X \Rightarrow Y}{\Gamma \Rightarrow X \supset Y}$

6 $(\neg)$ $\dfrac{\Gamma, X \Rightarrow \mathcal{F}}{\Gamma \Rightarrow \neg X}$

7 $(\equiv)$ $\dfrac{\Gamma, X \Rightarrow Y \quad \Gamma, Y \Rightarrow X}{\Gamma \Rightarrow X \equiv Y}$

8 $(\mathcal{F})$ $\dfrac{\Gamma \Rightarrow X, \Gamma \Rightarrow \neg X}{\Gamma \Rightarrow \mathcal{F}}$

*Règles d'élimination*

9 $(\wedge)$ $\dfrac{\Gamma \Rightarrow X \wedge Y}{\Gamma \Rightarrow X} \qquad \dfrac{\Gamma \Rightarrow X \wedge Y}{\Gamma \Rightarrow Y}$

10 $(\vee)$ $\dfrac{\Gamma \Rightarrow X \vee Y \quad \Gamma, X \Rightarrow Z \quad \Gamma, Y \Rightarrow Z}{\Gamma \Rightarrow Z}$

11 $(\supset)$ $\dfrac{\Gamma \Rightarrow X \supset Y}{\Gamma, X \Rightarrow Y}$

12 $(\neg)$ $\dfrac{\Gamma \Rightarrow X, \Gamma \Rightarrow \neg X}{\Gamma \Rightarrow \mathcal{F}} \qquad \dfrac{\Gamma \Rightarrow \neg\neg X}{\Gamma \Rightarrow X}$

13 $(\equiv)$ $\dfrac{\Gamma \Rightarrow X \equiv Y}{\Gamma, X \Rightarrow Y} \qquad \dfrac{\Gamma \Rightarrow X \equiv Y}{\Gamma, Y \Rightarrow X}$

14 $(\mathcal{F})$ $\dfrac{\Gamma \Rightarrow \mathcal{F}}{\Gamma \Rightarrow X}$

(La constante propositionnelle $\mathcal{F}$ dénote la valeur de vérité **F**.)

**Figure 3.32** *Règles de déduction naturelle dans le langage des séquents*

permet de *recouvrer* la dynamique de la démonstration effectuée en déduction naturelle en donnant une forme *déclarative* à ce qui était *procédural* dans l'exécution des preuves en déduction naturelle. Il ne s'agit pas de revenir à la déduction axiomatique et à ses aléas, mais d'ajouter aux facilités de la déduction naturelle une méthode transparente de gestion pour les prémisses auxiliaires ou non auxiliaires.

La *lecture* d'une démonstration construite par les méthodes de la déduction naturelle est pénible. Il faut souvent sauter plusieurs lignes pour retrouver l'hypothèse auxiliaire dont dérive une certaine formule. La déduction naturelle écrite dans le langage des séquents évite cet inconvénient. On conserve à chaque étape le stock des prémisses et on prélève dans celui-ci ce dont on a besoin au fur et à mesure que la démonstration progresse. (Scott et al, III & IV, 1981, p.139) En outre, le passage se fait continûment d'une ligne à la suivante dans chaque branche de l'arbre de preuve. Les feuilles à l'extrémité de chaque branche étant constituées par des "quasi-axiomes" de la forme $\overline{X \supset X}$. Le trait horizontal au-dessus de $X \supset X$ indique qu'il ne s'agit pas vraiment d'un axiome mais plutôt de la sortie d'une règle dont l'entrée est la classe vide des prémisses. On n'a donc pas quitté la déduction naturelle pour revenir subrepticement à la méthode axiomatique. On jugera des mérites de cette notation comme instrument de *communication* des preuves en comparant la lisibilité de la première démonstration que nous avons donnée du raisonnement 3.18 à celle de la démonstration exposée dans la figure 3.33. On dispose de l'ensemble des prémisses $\Gamma = \{s \supset (b \vee t), (b \vee a) \supset (r \wedge m), \neg r\}$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{\Gamma, b, (b \vee a) \supset (r \wedge m) \Rightarrow (b \vee a) \supset (r \wedge m)}}{\Gamma, b \Rightarrow (b \vee a) \supset (r \wedge m)}}{\Gamma, b, (b \vee a) \Rightarrow (r \wedge m)}}{\Gamma, b, (b \vee a) \Rightarrow r} \qquad \dfrac{\dfrac{\dfrac{}{\Gamma, b, (b \vee a), \neg r \Rightarrow \neg r}}{\Gamma, b, (b \vee a) \Rightarrow \neg r}}{\Gamma, b, (b \vee a) \Rightarrow \mathcal{F}}}{\Gamma, b \Rightarrow \neg(b \vee a)}$$

$$\dfrac{\dfrac{\dfrac{}{\Gamma, b \Rightarrow b}}{\Gamma, b \Rightarrow b \vee a} \qquad \dfrac{\dfrac{\dfrac{}{\Gamma, b \Rightarrow \mathcal{F}}}{\Gamma, b \Rightarrow t}}{\Gamma, s, b \Rightarrow t} \qquad \dfrac{\dfrac{}{\Gamma, t \Rightarrow t}}{\Gamma, s, t \Rightarrow t}}{\Gamma, s \Rightarrow t}$$

$$\dfrac{\dfrac{}{\Gamma \Rightarrow s \supset (b \vee t)}}{\Gamma, s \Rightarrow (b \vee t)} \qquad \dfrac{\Gamma, s \Rightarrow t}{\Gamma \Rightarrow s \supset t}$$

**Figure 3.33** *Dérivation de 3.18 en déduction naturelle formulée en séquents*

La *facilité de lecture* ne doit pas faire illusion. Cette facilité ne s'étend pas à la *découverte* de la preuve. Plusieurs règles sont *non déterministes* quand on les lit de haut en bas, c'est-à-dire en descendant de leur prémisse à leur conclusion. La *règle d'élimination* de $\mathcal{F}$ en est l'illustration la plus

claire puisqu'elle s'énonce "Si une contradiction découle de l'ensemble $\Gamma$, alors n'importe quelle proposition découle de $\Gamma$". Il en va de même de la *règle d'atténuation* qui permet d'ajouter n'importe quoi à l'ensemble des prémisses, ou encore de la *règle d'introduction* du $\vee$. Parmi les formules en nombre infini que l'on peut introduire à l'aide de cette règle, qu'on appelle parfois la règle de l'addition, il s'agit de repérer celle qui sera utile à l'objectif visé, à savoir démontrer $S \supset T$, exactement comme il fallait découvrir, parmi toutes les substitutions possibles, c'est-à-dire uniformes, celle qui est utile dans une démonstration axiomatique.

*Appendice.*

Le métathéorème de la déduction permet de déplacer le *turnstile* vers la gauche. Il faut, avant cela, l'introduire. A cette fin, on utilise les formes du *modus ponens* de la note 14 (p. 127) et de la figure 3.30 (p. 150) ainsi que la propriété de réflexivité et de monotonie.

Exemple d'introduction du *turnstile* ($\vdash$)

$$\frac{p, p \supset q \vdash q \text{ (M.P.)} \qquad q \supset r \vdash q \supset r \text{ (Réfl.)}}{p, p \supset q, q \supset r \vdash r} \text{ (M.P.)}$$

Deux formes de *modus ponens* ont été utilisées, ainsi que la réflexivité de la dérivabilité représentée par le *turnstile*. La deuxième est de la forme

$$\frac{\Gamma \vdash X \qquad \Delta \vdash X \supset Y}{\Gamma, \Delta \vdash Y} \text{ M.P.}$$

# Chapitre 4
# Le calcul des propositions – méthodes analytiques –

A l'exception de la méthode des tableaux matriciels, toutes les méthodes exposées dans le chapitre 3 nous imposaient, pour dériver une conclusion, de former une suite de formules dont certaines n'étaient ni des parties, ni des transformées de la conclusion, mais des formules indépendantes (axiomes, hypothèses, hypothèses auxiliaires). En d'autres termes, il fallait *ajouter* quelque formule exogène. C'est pour cette raison que nous avons appelé *synthétique* ces méthodes. Dans le chapitre 4, nous présenterons des méthodes que nous appelerons *analytiques*, parce que la conclusion est obtenue soit en *décomposant* [la négation de] la formule à démontrer en ses parties, soit en transformant la formule à démontrer sans y ajouter d'éléments extérieurs.

## 4.1. La notion de séquent

Soient $\alpha, \beta, \gamma, \delta$ des ensembles d'expressions bien formées :

$$\begin{aligned} \alpha &= \{X_1, \ldots, X_m\} \\ \beta &= \{Y_1, \ldots, Y_n\} \\ \gamma &= \{Z_1, \ldots, Z_l\} \\ \delta &= \{U_1, \ldots, U_k\}. \end{aligned}$$

Si nous voulions exprimer avec les symboles dont nous disposons actuellement que l'*une au moins* des expressions bien formées contenues dans $\gamma \cup \delta$ est dérivable de la *conjonction* des membres de $\alpha \cup \beta$, nous pourrions le faire à l'aide de la formule

(4.1) $\quad X_1 \wedge \cdots \wedge X_m \wedge Y_1 \wedge \cdots \wedge Y_n \vdash Z_1 \vee \cdots \vee Z_l \vee U_1 \vee \cdots \vee U_k,$

à condition que les ensembles $\alpha, \beta, \gamma, \delta$ ne soient pas infinis. Nos règles de formation, en effet, ne prévoient pas de conjonctions infinies ou de disjonctions infinies. Pour pouvoir traiter aussi les ensembles infinis et mélanger sans danger les *expressions bien formées* aux *ensembles d'expressions bien formées*[1], nous remplacerons les $\wedge$ et $\vee$ par des *virgules*, en nous rappelant que celles-ci ont une signification différente selon qu'elles sont à gauche ou à droite de $\vdash$ : à gauche, elles remplacent des $\wedge$, à droite, des $\vee$.

(4.2) $\quad X_1, \ldots, X_m, Y_1, \ldots, Y_n \vdash Z_1, \ldots, Z_l, U_1, \ldots, U_k$

Le prédicat métalinguistique "est dérivable de" n'est pas une notion primitive. C'est une notion définie : une formule $\Phi$ est dite dérivable d'un ensemble de formules $\Gamma$, ce qui se note

(4.3) $\quad \Gamma \vdash \Phi,$

si et seulement s'il existe une dérivation de $\Phi$ à partir de l'ensemble $\Gamma$. Rien n'empêche cependant de prendre $\vdash$ comme *terme primitif* et de lui conférer un sens grâce à un système d'axiomes et de règles d'inférence. Pour éviter toute confusion, on se donne un nouveau symbole $\Rightarrow$, qui remplacera $\vdash$ et sera traité comme primitif. La proposition métalinguistique 4.2 se réécrira

(4.4) $\quad \langle X_1, \ldots, X_m, Y_1, \ldots, Y_n \rangle \Rightarrow \langle Z_1, \ldots, Z_l, U_1, \ldots, U_k \rangle.$

**Définition 4.1 : le séquent.** On appelle *séquent* une expression de la forme[2]

(4.5) $\quad \langle \alpha, \beta, X \ldots \rangle \Rightarrow \langle \gamma, \delta, Y \ldots \rangle,$

où l'expression à gauche du signe $\Rightarrow$ s'appelle *antécédent*, et l'expression à droite s'appelle *succédent*. $X$ et $Y$ sont des formules, $\alpha, \beta, \gamma, \delta$ sont des ensembles de formules. Rappelons que l'on peut *mélanger* les formules et les ensembles de formules grâce à un emploi approprié des virgules.

On a aussi les cas dégénérés suivants :

- Les séquents dont l'antécédent est vide :

(4.6) $\quad \Rightarrow \langle \gamma, \ldots, \delta \ldots \rangle$

- Les séquents dont le succédent est vide :

(4.7) $\quad \langle \alpha, \ldots, \beta \ldots \rangle \Rightarrow$

- Le séquent dont l'antécédent et le succédent sont vides :

(4.8) $\quad \Rightarrow$

---

[1] comme on le fait dans la formulation du métathéorème de la déduction :

$$\frac{\Gamma, X \vdash Y}{\Gamma \vdash X \supset Y}.$$

[2] Cette expression se lit : "De la conjonction des expressions bien formées qui sont membres de $\alpha, \beta \ldots$ découle la disjonction des expressions bien formées qui sont membres de $\gamma, \delta \ldots$"

### 4.1.1. *Antécédents conjoints, succédents disjoints*

Pourquoi fait-on un sort à la relation de dérivabilité existant entre une *conjonction* d'expressions bien formées et une *disjonction* d'expressions bien formées ? On trouve dans Scott et al. (1981, tome 2, p. 138) une réponse qui est, en substance, celle-ci : de même que toute *formule* $\Phi$ inconsistante possède un dual, à savoir $\neg\Phi$, qui est une formule valide, ainsi à tout *ensemble* inconsistant d'expressions bien formées correspond un séquent qu'on pourrait, par abus de langage, appeler son dual. L'exemple suivant montre comment on obtient un séquent à partir du complément d'un ensemble inconsistant. Rappelons que $\overline{r}$ et $\neg r$ sont des variantes notationnelles.

Soit l'ensemble inconsistant

(4.9) $\{p \vee q, p \supset r, q \supset s, \overline{r}, \overline{s}\}$.

Transformons-le en conjonction :

(4.10) $(p \vee q) \wedge (p \supset r) \wedge (q \supset s) \wedge \overline{r} \wedge \overline{s}$.

Nions cette conjonction. Nous obtenons la forme propositionnelle valide

(4.11) $\neg((p \vee q) \wedge (p \supset r) \wedge (q \supset s) \wedge \overline{r} \wedge \overline{s})$,

qui est équivalente à d'autres formes propositionnelles et notamment à

(4.12) $(((p \vee q) \wedge (p \supset r) \wedge (q \supset s)) \supset (r \vee s))$,

qui a la forme

(4.13) $(X_1 \wedge \cdots \wedge X_m) \supset (Y_1 \vee \cdots \vee Y_n)$,

et qui correspond au séquent

(4.14) $X_1, \ldots, X_m \Rightarrow Y_1, \ldots, Y_n$ où $m = 3$ et $n = 1$.

*Remarque.* Notons que le symbole $\Rightarrow$ n'est pas un connecteur mais un prédicat métalinguistique. Il n'est pas itérable. Alors que les *formes* propositionnelles du type

(4.15) $X \supset (Y \supset Z)$

sont des expressions bien formées, un séquent du type

(4.16) $X \Rightarrow (Y \Rightarrow Z)$

est mal formé.

### 4.1.2. *Le calcul des séquents*

Gentzen a formulé un système axiomatique dans lequel le symbole $\Rightarrow$ est pris comme symbole primitif. Ce système s'appelle *calcul des séquents*.

### 4.1.2.1. *Axiomes*

Un séquent est un axiome si et seulement si son antécédent et son succédent ont en commun au moins une proposition atomique. Tous les axiomes sont valides. En effet, pour rendre faux un conditionnel dont l'antécédent est une conjonction et le conséquent une disjonction, c'est-à-dire un énoncé de la forme

(4.17) $(X_1 \wedge \cdots \wedge p \wedge \cdots \wedge X_n) \supset (Y_1 \vee \cdots \vee p \vee \cdots \vee Y_m)$,

il faut rendre vrais tous les membres de l'antécédent, et faux tous ceux du conséquent. Cela exigerait que l'on donne aux deux occurrences de l'atome propositionnel $p$ une valeur différente, ce qui serait une infraction à la règle d'assignation de valeur aux variables, assignation qui doit être uniforme (l'assignation est une *application* cf. 7.3.1) .

### 4.1.2.2. *Règles structurales*

Comme l'antécédent et le succédent sont des *suites*, donc des n-uples ordonnés, on ne peut pas permuter librement les éléments ou supprimer un élément qui se répète, comme on pourrait le faire si l'antécédent et le succédent étaient des ensembles (non ordonnés). Or on voudrait pouvoir permuter les expressions bien formées dans un même membre du séquent (c'est-à-dire dans l'antécédent ou dans le succédent, en vertu de la commutativité du $\wedge$ et du $\vee$). De même, on voudrait pouvoir contracter $X, X$ en $X$ dans le *succédent* ou ventiler $X$ en $X, X$ dans l'*antécédent*, en vertu de la loi d'idempotence pour $\vee$ et pour $\wedge$.

Enfin, sachant que l'antécédent est une conjonction, on voudrait pouvoir y ajouter impunément des prémisses inutiles, en vertu de la *monotonie* de la relation de conséquence logique, et on voudrait pouvoir ajouter une proposition quelconque au succédent en vertu de la règle de l'addition (Si $Y$ est dérivable de $X$, alors $Y \vee Z$ l'est aussi).

Gentzen formule des règles d'inférence qui autorisent d'exécuter les opérations précitées sur les suites.

Dans un séquent, il y a ce qu'on appelle la *formule principale* — c'est la formule sur laquelle on travaille — et les autres formules, auxquelles on ne touche pas. Ces dernières seront collectivement désignées par $\alpha$ quand elles sont dans l'antécédent et par $\beta$ quand elles sont dans le succédent. On appelle parfois $\alpha$ et $\beta$ du nom de *formules additionnelles*. La répartition des formules en formule principale et formules additionnelles change au cours du processus de démonstration.

*Règle de permutation (dans l'antécédent et le succédent)*

$$\frac{\alpha \Rightarrow \beta, X}{\alpha \Rightarrow X, \beta} \qquad \frac{X, \alpha \Rightarrow \beta}{\alpha, X \Rightarrow \beta}$$

*Règle de contraction (dans l'antécédent et le succédent)*

$$\frac{\alpha \Rightarrow \beta, X, X}{\alpha \Rightarrow \beta, X} \qquad \frac{X, X, \alpha \Rightarrow \beta}{X, \alpha \Rightarrow \beta}$$

*Règle d'atténuation (dans l'antécédent et le succédent)*

$$\frac{\alpha \Rightarrow \beta}{\alpha \Rightarrow \beta, X} \qquad \frac{\alpha \Rightarrow \beta}{X, \alpha \Rightarrow \beta}$$

*Remarque.* Pour la logique classique, on peut se passer de ces règles sans appauvrir le système, à condition de remplacer les *suites* figurant dans les rôles d'antécédent et de succédent par des *ensembles*. Nous adopterons cette présentation simplifiée.

4.1.2.3. *Règles d'introduction des connecteurs propositionnels*

Sachant que le $\wedge$ et le $\vee$ sont associatifs, on accepte les conjonctions et les disjonctions à $n$ éléments $(X_1 \wedge \cdots \wedge X_n)$ et $(X_1 \vee \cdots \vee X_n)$, que l'on écrira de manière plus concise $\bigwedge_i X_i$ et $\bigvee_i X_i$, si $n$ est connu.

(Les deux règles * de la figure 4.1 introduisent des ramifications qui peuvent ne pas être binaires ($n$ peut être plus grand que 2, mais doit être fini).)

4.1.2.4. *Règle d'élimination*

$$\frac{\alpha \Rightarrow \beta, X \qquad X, \gamma \Rightarrow \delta}{\alpha, \gamma \Rightarrow \beta, \delta} \text{ coupure}$$

$$\neg \Rightarrow \frac{\alpha \Rightarrow \beta, X}{\neg X, \alpha \Rightarrow \beta} \qquad \frac{X, \alpha \Rightarrow \beta}{\alpha \Rightarrow \beta, \neg X} \Rightarrow \neg$$

$$\supset \Rightarrow \frac{\alpha \Rightarrow \beta, X \quad Y, \alpha \Rightarrow \beta}{X \supset Y, \alpha \Rightarrow \beta} \qquad \frac{X, \alpha \Rightarrow \beta, Y}{\alpha \Rightarrow \beta, X \supset Y} \Rightarrow \supset$$

$$\equiv \Rightarrow \frac{X, Y, \alpha \Rightarrow \beta \quad \alpha \Rightarrow \beta, X, Y}{X \equiv Y, \alpha \Rightarrow \beta} \qquad \frac{X, \alpha \Rightarrow \beta, Y \quad Y, \alpha \Rightarrow \beta, X}{\alpha \Rightarrow \beta, X \equiv Y} \Rightarrow \equiv$$

$$\wedge \Rightarrow \frac{X_1, \ldots, X_n, \alpha \Rightarrow \beta}{(X_1 \wedge \cdots \wedge X_n), \alpha \Rightarrow \beta} \qquad \frac{\alpha \Rightarrow \beta, X_1 \quad \ldots \quad \alpha \Rightarrow \beta, X_n}{\alpha \Rightarrow \beta, (X_1 \wedge \cdots \wedge X_n)} \Rightarrow \wedge^*$$

$$\vee^* \Rightarrow \frac{X_1, \alpha \Rightarrow \beta \quad \ldots \quad X_n, \alpha \Rightarrow \beta}{(X_1 \vee \cdots \vee X_n), \alpha \Rightarrow \beta} \qquad \frac{\alpha \Rightarrow \beta, X_1, \ldots, X_n}{\alpha \Rightarrow \beta, (X_1 \vee \cdots \vee X_n)} \Rightarrow \vee$$

**Figure 4.1** *Liste des règles d'introduction de connecteurs*

### 4.1.3. *La démonstration dans le calcul des séquents*

**Définition 4.2 : la démonstration en calcul des séquents.** Une démonstration du calcul des séquents est un arbre constitué d'une ou plusieurs successions de séquents qui aboutissent toutes à un même séquent, le *séquent final*, situé à la racine [adapté de Ladrière, 1951, p. 359].

Chaque séquent situé à l'extrémité supérieure d'une branche est un axiome et chacun des autres séquents est, soit l'un des séquents supérieurs, soit le séquent inférieur, d'une des règles d'inférence. De plus, un séquent ne peut être séquent supérieur que d'une seule règle. C'est cette dernière contrainte qui impose à la démonstration en calcul des séquents la forme d'un arbre généalogique (Ladrière, 1951, *Ibid.*).

### 4.1.4. *Réversibilité des règles*

Les règles d'inférence doivent, pour être correctes, transporter la valeur distinguée, c'est-à-dire la *vérité* des propositions ou la *validité* des formes propositionnelles. On n'exige pas qu'elles soient, par surcroît, bons conducteurs du faux ou du non-valide.

La conclusion d'un *modus ponens*, pris comme schéma d'inférence du langage-objet ($p, p \supset q \vdash q$) par exemple, est une proposition vraie *si*, mais pas *seulement si*, les prémisses sont vraies. La conclusion d'un *modus ponens*, pris comme règle d'inférence du métalangage ($X, X \supset Y \vdash Y$), est une forme propositionnelle valide *si*, mais pas *seulement si*, les prémisses sont des formes propositionnelles valides.

On trouve sans peine des inférences de la langue naturelle de la forme :

(4.18) $$\frac{p, p \supset q}{q}.$$

Par exemple,

$$\text{(4.19)} \quad \frac{2+2=5,\ \text{si}\ 2+2=5\ \text{alors}\ 2+2=4\ \text{ou}\ 2+2=5}{(2+2=4)\ \text{ou}\ (2+2=5)},$$

où la conclusion est vraie, mais où l'une des prémisses est fausse.

Et, pour illustrer le deuxième cas, si on substitue $p$ à $X$ et $p \supset (p \vee q)$ à $Y$ dans la règle

$$\text{(4.20)} \quad \frac{X, X \supset Y}{Y},$$

on obtient

$$\text{(4.21)} \quad \frac{p, p \supset [p \supset (p \vee q)]}{p \supset (p \vee q)},$$

qui est une instance du *modus ponens* dans laquelle la conclusion est valide *sans* que toutes les prémisses le soient.

Dans la logique de la déduction naturelle, la règle d'*introduction* de la disjonction — la règle de l'addition — n'est manifestement pas réversible non plus :

$$\text{(4.22)} \quad \frac{X}{X \vee Y}.$$

Elle transmet la vérité ou la validité de la prémisse à la conclusion, mais ne transmet pas la vérité ou la validité de la conclusion aux prémisses. De "Il est chez lui ou il est sorti", on ne peut inférer validement "Il est chez lui". Dans le calcul des séquents privé de la coupure, au contraire, *toutes* les règles d'*introduction* sans exception sont réversibles, y compris celle de l'*introduction* de la disjonction [3].

Le calcul des séquents satisfait la condition suivante :

1. le séquent-conclusion est vrai relativement à une interprétation $\mathcal{I}$ *si et seulement si* tous les séquents-prémisses sont vrais. Cette condition est trivialement équivalente à celle-ci :
2. le séquent-conclusion est faux relativement à une interprétation $\mathcal{I}$ *si et seulement si* au moins un des séquents-prémisses est faux.

Pour démontrer la réversibilité des règles d'introduction, il suffit d'examiner chacune des dix règles d'introduction de connecteur et de donner aux constituants du séquent-conclusion des valeurs qui rendent celui-ci faux. On s'aperçoit alors qu'en vertu du sens des connecteurs, l'un au moins des séquents-prémisses (ou le séquent-prémisse quand il n'y en a qu'un) est faux. Puis on fait le travail inverse. On donne une valeur aux constituants des séquents-prémisses qui rend au moins l'un d'eux faux, et on s'aperçoit que le séquent-conclusion l'est aussi. On doit pour cela se souvenir qu'une évaluation $\mathcal{I}$ rend le séquent

[3] Rappelons que la règle de coupure, qui n'est pas réversible, n'est pas une règle d'*introduction*, mais bien une règle d'élimination.

$X_1, \ldots, X_n \Rightarrow Y_1, \ldots, Y_m$ vrai si et seulement si elle rend vrai le conditionnel $X_1 \wedge \cdots \wedge X_n \supset Y_1 \vee \cdots \vee Y_m$.

### 4.1.5. *Démonstration de la réversibilité d'une règle*

Nous illustrerons cette méthode en l'appliquant à la règle d'*introduction* de la *disjonction* à gauche. Sous sa forme la plus générale, la règle s'énonce ainsi :

$$(4.23) \qquad \frac{X_1, \alpha \Rightarrow \beta \quad \ldots \quad X_n, \alpha \Rightarrow \beta}{(X_1 \vee \cdots \vee X_n), \alpha \Rightarrow \beta}.$$

Prenons le cas où $n = 2$. La règle a la forme de l'arbre à deux branches de la figure 4.2.

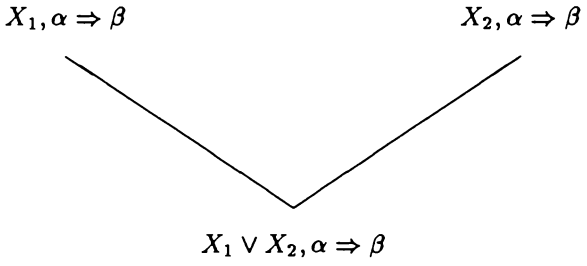



**Figure 4.2** *Introduction d'une disjonction à deux membres*

On se souvient que la lettre auxiliaire $\alpha$ représente la partie de l'antécédent du séquent sur laquelle on ne travaille pas, à savoir une *conjonction* de formules $(A_1 \wedge \cdots \wedge A_m)$, et que $\beta$ représente la partie *succédent* du séquent sur lequel on ne travaille pas, à savoir une *disjonction* de formules $(B_1 \vee \cdots \vee B_k)$:

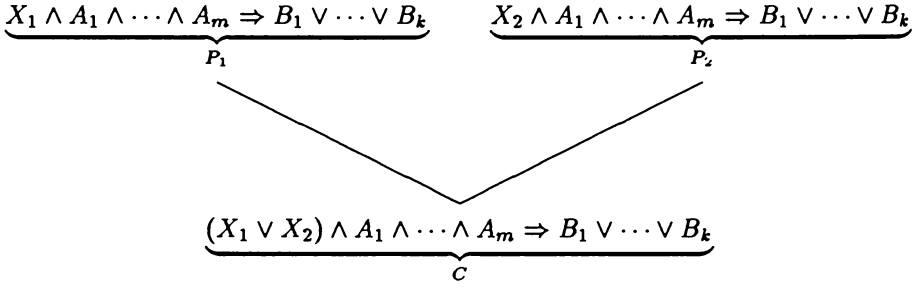



**Figure 4.3** *Développement des formules auxiliaires*

Pour rendre fausse la conclusion $X_1 \vee X_2, \alpha \Rightarrow \beta$, qui est une abréviation de $(X_1 \vee X_2) \wedge A_1 \wedge \cdots \wedge A_m \Rightarrow B_1 \vee \cdots \vee B_k$, il faut rendre vraies toutes les formules

de l'antécédent et fausses toutes celles du succédent. Il faut, en particulier, rendre vraie la formule $(X_1 \vee X_2)$, ce qui exige — en vertu du sens de $\vee$ — que soit $X_1$, soit $X_2$, soit $X_1$ et $X_2$ soient vrais. Si l'une de ces trois éventualités est réalisée, alors l'une au moins des prémisses est fausse (celle qui contient $X_i$ vrai). La réciproque est immédiate : si une des prémisses est fausse, un des $X$ est vrai (l'$X_i$ qui figure dans la prémisse fausse). Dans cette interprétation, la conclusion sera fausse aussi car tous les membres de l'antécédent (y compris $(X_1 \vee X_2)$) seront vrais et tous les membres du succédent seront faux.

### 4.1.6. *Méthode pour démontrer un séquent*

Comme nous l'avons vu plus haut, il résulte de la réversibilité des règles du calcul des séquents qu'un schéma à la racine d'un arbre de preuve est valide *si et seulement si* toutes les feuilles aux extrémités des branches sont des axiomes, et qu'il est falsifiable *si et seulement si* au moins une des feuilles à l'extrémité d'une branche est falsifiable. De là découle une méthode de construction de preuve qui, au moins pour ce qui est du calcul des propositions, et pour autant que l'on se passe de la règle de coupure, est *mécanique*. On peut, en effet, partir du séquent à démontrer, essayer de le falsifier, et constater qu'on échoue dans cette tentative, ce qui constitue une preuve indirecte du séquent . Il faut, pour cela, que l'on applique les règles d'introduction à *rebours*, c'est-à-dire comme des règles d'élimination et que, ce faisant, on arrive sur chaque branche à un nœud étiqueté par un axiome. Il suffira de relire ensuite la démonstration dans l'autre sens — en allant cette fois des axiomes-séquents à la conclusion — pour obtenir une *preuve directe*.

En partant du séquent à démontrer (que l'on met à droite de la flèche) et en appliquant les règles d'introduction à l'envers, on arrive soit à une preuve du séquent, soit à la construction d'un contre-exemple. On engendrera autant de contre-exemples, c'est-à-dire de situations qui falsifient le séquent, que l'on parcourra de branches à l'extrémité desquelles se trouve un nœud terminal qui n'est pas un axiome. Un nœud est terminal sans être un axiome s'il est étiqueté par un séquent formé d'une flèche et uniquement de propositions atomiques dont aucune ne figure à la fois à gauche et à droite du séquent.

### 4.1.7. *Exemple de démonstration de séquent*

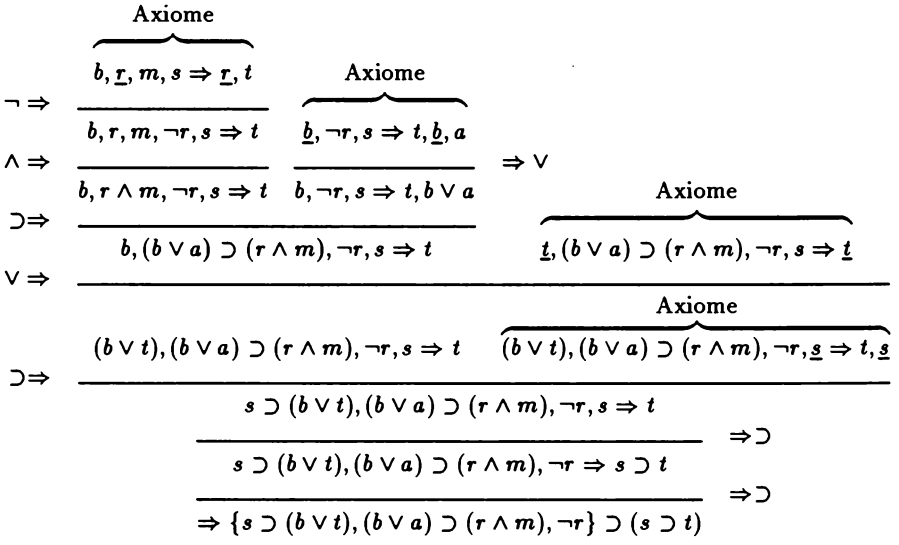



**Figure 4.4** *Dérivation de 3.18 en calcul des séquents*

### 4.1.8. *La part du choix dans la preuve en calcul des séquents*

A la deuxième ligne en montant dans l'arbre, nous étions confrontés à un choix entre quatre possibilités : appliquer une règle d'élimination au premier ou au second conditionnel figurant dans l'antécédent du séquent, appliquer une telle règle à la négation ou éliminer le conditionnel figurant dans le succédent. On peut montrer que ces choix ont une incidence sur la *longueur* des démonstrations mais non sur leur *issue*, contrairement au choix des points de départ en déduction axiomatique.

Une fois la décision prise de travailler sur l'une des quatre formules disponibles, la nature du connecteur principal de la formule traitée dicte le choix de la règle à appliquer et l'application de la règle elle-même ne nécessite aucun choix. Cet état de choses contraste avec la situation à laquelle nous étions confrontés en déduction axiomatique ou en déduction naturelle.

Dans une déduction axiomatique, on doit faire des substitutions sur certaines variables des axiomes ou des théorèmes qu'on prend comme point de départ. Le choix d'une substitution qui soit non seulement correcte — c'est-à-dire uniforme — mais de surcroît utile, exige un effort d'imagination. Les substitutions

ne se révèlent utiles que *rétrospectivement*. On ne peut donc pas trouver la substitution à faire par simple inspection de la formule canonique de la thèse dont on part. En revanche, on peut recouvrer, après coup, la forme canonique à partir de la forme instanciée (qui en est déduite par substitution). Par exemple, la simple inspection de la formule

$$(4.24)\quad ((p \supset q) \supset (p \supset \neg\neg q)) \supset (((p \supset \neg\neg q) \supset (\neg q \supset \neg p)) \supset ((p \supset q) \supset (\neg q \supset \neg p)))$$

permet d'y voir une instance du théorème

$$(4.25)\quad (p \supset q) \supset ((q \supset r) \supset (p \supset r)).$$

Inversement, lorsqu'on lit la *conclusion* d'une application de la règle du *modus ponens*, on ne peut recouvrer les prémisses dont la conclusion est déduite. Au contraire, si on lit les *prémisses*, par exemple, $(p \vee q) \supset (r \vee s)$ et $(p \vee q)$, il n'y a qu'une seule conclusion possible par *modus ponens*, à savoir $(r \vee s)$. Appelant le passage de l'axiome sous sa forme canonique à l'axiome sous sa forme instanciée ou le passage des prémisses d'une règle à sa conclusion, le mode "synthétique", et le mouvement inverse, le "mode du raffinement", Wallen (1987) décrit cette asymétrie en ces termes : les *axiomes* lus sur le mode synthétique sont fortement "non déterministes", ils laissent beaucoup de choix, alors qu'ils sont déterministes quand ils sont lus selon le "mode du raffinement". C'est l'inverse pour la règle du *modus ponens*.

En travaillant avec des règles réversibles uniquement, le calcul des séquents extirpe toute trace d'indéterminisme au niveau *local* de l'application des règles : une fois que l'on a choisi la formule sur laquelle on va travailler, les règles déterminent entièrement les transformations à effectuer (sauf dans le cas de la règle de coupure). Le caractère *mécanique* des règles utilisées est une condition nécessaire, mais non suffisante, pour qu'on puisse obtenir une procédure de décision. Pour que l'on puisse ériger le calcul propositionnel des séquents en méthode de décision, il faut, en outre, être assuré que le processus mécanique *s'arrête* après un nombre fini de pas. Or cette seconde exigence est satisfaite. Une autre particularité du calcul des séquents contribue à satisfaire cette exigence : le fait que les preuves de ce calcul faites sans recourir à la règle de coupure satisfont une condition à laquelle on a donné le nom de principe de la sous-formule.

### 4.1.9. *Le calcul des séquents avec coupure*

Les règles d'inférence de la déduction naturelle et du calcul des séquents *avec coupure* se répartissent en deux groupes : les règles d'*introduction* et les règles d'*élimination*. Dans le calcul des séquents, il ne subsiste qu'une seule règle d'élimination : la coupure. Gentzen a démontré le théorème fondamental suivant :

(4.26) Toute démonstration d'une expression bien formée dans le calcul des séquents peut se transformer en une démonstration qui a même formule finale et dans laquelle aucune application de la règle de coupure n'est faite.

Cette démonstration est difficile. Elle exige que l'on fasse deux raisonnements par induction (mathématique) intégrés l'un à l'autre. On en trouvera une présentation claire dans Ladrière, 1951, pp. 360–365.

Le calcul des séquents avec coupure apparaît, *rétrospectivement*, comme un *calcul intermédiaire* entre la déduction naturelle, qui contient *autant* de règles d'élimination *que* de règles d'introduction, et le calcul des séquents sans coupure qui ne comporte *aucune* règle d'élimination et qui possède la propriété de sous-formule.

### 4.1.10. *Le principe de la sous-formule*

Les démonstrations construites dans le calcul des séquents sans coupure possèdent une propriété remarquable qui fait défaut aux démonstrations en général : elles satisfont le principe de la sous-formule. Ce principe s'énonce comme suit : "... chaque formule qui est membre d'un séquent quelque part dans la démonstration est un constituant (une sous-formule) *dans la conclusion de la démonstration*" (Robinson, 1979, pp. 108–109).

Ce n'est pas le cas dans la démonstration axiomatique. Dans la démonstration de la formule $(p \supset (p \supset (p \supset p))$ obtenue à partir de $((q \supset p) \supset r) \supset (p \supset r)$, de $p \supset (q \supset p)$ et d'autres théorèmes, il est clair que la formule citée en dernier lieu n'est pas une sous-formule du théorème à démontrer.

Le principe de la sous-formule a une conséquence théorique et pratique immédiate. Pour le calcul propositionnel formalisé dans le calcul des séquents, la recherche systématique d'un contre-exemple à un séquent fini aboutira, au bout d'un nombre fini de pas, soit à la production effective d'un contre-exemple, soit à la démonstration qu'il n'en existe pas, toutes les voies de recherche se fermant sur des axiomes. La recherche d'un contre-exemple se fait, en effet, par l'application des règles d'introduction lues à l'envers, lesquelles font *décroître à chaque pas la complexité des formules manipulées*, et remonter de proche en proche jusqu'à des "ancêtres" atomiques. Le principe de la sous-formule nous assure que, au moins pour les séquents finis, c'est-à-dire formés d'ensembles finis de formules, la quête d'un contre-exemple se *terminera*.

### 4.1.11. *Complétude du calcul propositionnel des séquents*

On a vu que dans le calcul propositionnel des séquents, dont les règles ont été formulées en 4.1.2, il n'y a que deux issues possibles : ou la recherche d'un contre-exemple à la formule $X$ réussit, ce qui signifie que $\neg X$ est

sémantiquement consistant (*satisfiable* en anglais), ou une *preuve* de $X$, à partir d'axiomes, est trouvée.

Formellement, cette alternative s'écrit :

(4.27) $\neg X$ est sémantiquement consistant ou $X$ est démontrable.

Or en vertu de l'équivalence $p \vee q \equiv \neg p \supset q$, cela est équivalent à

(4.28) Si $\neg X$ est sémantiquement inconsistant, alors $X$ est démontrable,

et, comme l'inconsistance de $\neg X$ est identique à la validité de $X$, on a

(4.29) Si $X$ est valide, alors $X$ est démontrable,

qui est la complétude recherchée.

### 4.1.12. *La portée philosophique du théorème de Gentzen*

On notera le contraste entre les démonstrations opérées dans le cadre du calcul des séquents sans coupure, qui sont *entièrement analytiques* et d'où sont éliminés tous les *détours*, et les *moyens termes* et la première espèce de démonstration introduite par le fondateur de la logique (Aristote), à savoir le syllogisme, qui est précisément une forme de coupure. Gentzen a véritablement conféré un sens nouveau à la notion de *démonstration* :

> ... celle-ci ne consiste plus à comparer certains termes de façon à dégager certaines liaisons ni à passer d'un certain ensemble de formules à une formule nouvelle au moyen de certaines règles de transformation, mais elle consiste à construire progressivement la formule à démontrer, élément par élément (Ladrière, 1951, p. 378).
>
> [Elle] ne procèdera plus du plus connu au moins connu, par propagation de l'évidence, comme dans une axiomatique à contenu intuitif. Elle ne passera plus de formules considérées arbitrairement comme valables à des formules nouvelles, comme dans une axiomatique de type formel : ... elle n'est plus passage, cheminement, liaison réglée, mais procédé de construction, loi de formation (Ladrière, 1951, p. 379).

Rappelons que la construction se fait à l'aide de matériaux prélevés dans la formule à démontrer que l'on a commencé par démembrer.

On arrive ainsi à la notion de "démonstrations canoniques", ce qui, note encore Ladrière, "réduit la part d'imprévisibilité dans le calcul". On doit cependant reconnaître que

> cette forme canonique n'est ... déterminée univoquement qu'à l'ordre près des schémas utilisés... (Ladrière, 1951, p. 380).

En d'autres termes, une démonstration par le calcul des séquents sans coupure est une *démonstration sans détour*, mais il peut exister plusieurs démonstrations sans détour qui sont de longueur inégale. Il reste donc un progrès à faire : trouver la démonstration *sans détour la plus courte*. C'est ce

que réalise la méthode des connexions de Bibel et la méthode des matrices de Wallen (la seconde sera exposée en 4.2).

### 4.1.13. *Calcul des séquents et déduction naturelle*

Au début de ce chapitre, nous avons opposé les méthodes ***synthétiques*** de preuve déductive aux méthodes ***analytiques***. Il importe de préciser que les bonnes propriétés de la méthode des séquents sans coupure n'appartiennent pas exclusivement à cette méthode. Le théorème de la forme normale de Prawitz (3.6.1) étend, contre toute attente, ces bonnes propriétés aux dérivations en déduction naturelle "normalisées", c'est-à-dire à une des méthodes que nous avons appelées synthétiques.

Quand la dérivation d'une formule $Y$ à partir des suppositions $X_1, \ldots, X_n$ a été ***normalisée***, chaque formule qui apparaît dans la dérivation doit être une ***sous-formule*** de $Y$ ou de l'une des $X_i$. D'autre part, l'application systématique du principe d'inversion permet d'éviter les détours. La normalisation en déduction naturelle est l'équivalent de l'élimination de la coupure dans le calcul des séquents. Ce qui reste propre au calcul des séquents, c'est — si l'on se borne au calcul des séquents sans coupure — le fait de posséder ces bonnes propriétés à l'origine, plutôt qu'à la suite de transformations comme celles qui normalisent un dérivation en déduction naturelle.

### 4.1.14. *Les constantes logiques dans le calcul des séquents*

Les auteurs qui tentent de définir les constantes logiques à partir des règles de déduction naturelle ont généralement une attitude discriminatoire à l'égard de certaines règles. Ils font porter le fardeau de la définition aux seules règles d'introduction. Parfois ils invoquent, tantôt les règles d'introduction pour tous les connecteurs sauf un, et les règles d'élimination pour ce dernier. C'est ce que fait Tennant qui propose dans *Anti-realism and Logic* (1986) une ingénieuse reconstruction rationnelle de l'acquisition du sens des connecteurs. Cette attitude discriminatoire est bien difficile à justifier quand on se place dans le contexte de la déduction naturelle. Elle est plus difficile à justifier encore quand on se place dans le cadre du calcul des séquents, puisque dans ce dernier il n'y a plus que des règles d'introduction (introduction à gauche et à droite, les premières correspondant aux règles d'élimination en déduction naturelle). Faute d'avoir rencontré une justification convaincante de l'attitude discriminatoire en question, nous adoptons la position de M. Kremer, qui

> [...] prend les règles d'introduction à gauche et à droite (dans le calcul des séquents) comme conférant ensemble (*jointly*) la signification des symboles introduits, et qui s'abstient de poser un ordre de priorité entre elles (Kremer, 1988, p. 54).

A l'intérieur du calcul des séquents, M. Kremer a réussi à formuler de manière rigoureuse les conditions qui doivent être satisfaites pour que la définition des constantes logiques fournies implicitement par les règles de calcul soit *non-créative*. Les deux conditions suivantes doivent être satisfaites.

1. La règle de coupure doit être éliminable. Rappelons que la règle de coupure est une règle dérivée qui a la forme suivante :

$$\frac{\alpha \Rightarrow \beta, X \quad X, \gamma \Rightarrow \delta}{\alpha, \gamma \Rightarrow \beta, \delta}$$

   C'est la seule règle du calcul des séquents qui élimine quelque chose (la formule $X$) et qui, de ce fait, mette en péril le principe de la sous-formule.

2. Aucune application des règles d'introduction pour les anciens symboles ne doit aboutir à l'*élimination* du symbole nouvellement introduit.

Pour comprendre la raison d'être de ces deux exigences, considérons une preuve qui nécessite d'appliquer les règles d'introduction d'un nouveau symbole. Si les deux exigences mentionnées plus haut sont satisfaites,

> [...] le nouveau symbole apparaîtra dans tout séquent postérieur dans la preuve et en particulier dans la conclusion. Les règles d'introduction du nouveau symbole ne permettront donc pas de prouver des séquents contenant seulement l'ancien lexique logique, mais qu'il aurait été impossible de prouver avec l'ancienne théorie [celle qui ne comportait pas le nouveau symbole] (M. Kremer, *Ibid.*, p. 63).

En d'autres termes, en nous conformant à ces deux exigences, nous forçons la théorie $T_{n+1}$, obtenue à partir de la théorie $T_n$ par l'adjonction de règles introduisant un nouveau symbole, à être une *extension conservatrice*[4] de la théorie $T_n$ et nous forçons du même coup les règles d'introduction à se comporter en *définitions non créatives* (cf. 1.6.5) au sens strict de ces termes.

La contrainte de non-créativité[5] imposée par Lesniewski aux définitions avait pour but de nous assurer, comme le note M. Kremer, que la définition proposée ne nous donne "rien d'autre que" la signification du symbole. Le critère d'éliminabilité est destiné à nous assurer que ladite définition nous donne "l'intégralité" de la signification du nouveau symbole. Si la définition des symboles logiques par les règles d'inférence du calcul des séquents satisfait

[4] La théorie $T_{n+1}$ est appelée extension conservatrice de la théorie $T_n$ si et seulement si tous les théorèmes de $T_{n+1}$ qui ne contiennent que des constantes de $T_n$ sont déjà des théorèmes de $T_n$ (Rogers, 1971, p. 100).

[5] Suppes définit ainsi la non-créativité :

> Une formule $S$ introduisant un nouveau symbole d'une théorie satisfait le critère de non-créativité si et seulement s'il n'y a pas de formule $T$, dans laquelle de nouveau symbole n'apparaît pas, telle que $S \supset T$ soit dérivable des axiomes et des définitions antérieures de la théorie alors que $T$ n'en est pas dérivable (Suppes, 1957, p. 154).

seulement la première exigence, elle n'est qu'une définition *partielle*, mais, dans la mesure où elle satisfait au moins la première exigence, elle est une définition au sens propre et non dans un sens analogique.

Tant que l'on se confine à la définition axiomatique, à la manière de Whitehead et Russell, on peut bien dire que les symboles logiques primitifs sont *définis* par les axiomes, mais dans ce contexte, on veut seulement dire que le sens des termes primitifs est *restreint* par l'obligation de donner un sens aux termes primitifs grâce auxquels les axiomes deviennent vrais.

Les axiomes dans lesquels figurent les termes primitifs sont, par ailleurs, des axiomes créatifs au même titre que les axiomes de l'arithmétique. Quand, au contraire, on se hisse au niveau du métalangage, niveau auquel sont formulées les règles de la déduction naturelle ou du calcul des séquents, il devient possible de dire que les règles sont des définitions contextuelles des constantes logiques dans un sens plus fort du mot "définition", à savoir dans le sens où une définition diffère d'un axiome par son caractère *non créatif*. Nous trouvons ici des arguments indépendants en faveur de la conception de la logique défendue par Etchemendy (voir 7.8.1), la conception de la logique comme métathéorie de l'inférence, plutôt que comme une théorie d'un type particulier de vérité qui serait en continuité avec l'arithmétique ou la géométrie.

### 4.1.15. *La méthode des tableaux sémantiques*

Au cours des années cinquante, Beth et Hintikka, indépendamment l'un de l'autre, ont inventé une nouvelle méthode de preuve, la méthode des tableaux sémantiques. Il ne s'agit nullement d'une variante notationnelle de la méthode des tables de vérité mais, au contraire, d'une méthode reposant sur des principes totalement différents.

Dans la *méthode des tableaux matriciels*, et donc aussi dans l'algorithme d'analyse de vérité de Quine, on assigne une valeur de vérité à des constituants indépendants. Puis on calcule la valeur de vérité du composé à partir de la valeur de vérité de ses constituants, la valeur de vérité du tout étant fonction de la valeur de vérité des parties. On va donc de l'*intérieur* vers l'*extérieur*.

Dans la *méthode des tableaux sémantiques*, au contraire, on fait la démarche inverse : on part de la conjecture que la formule que l'on a en face de soi peut être rendue fausse pour une certaine interprétation et on tire les conséquences de cette hypothèse pour ce qui est de la valeur de vérité des constituants : on va donc de l'*extérieur* vers l'*intérieur*, c'est-à-dire que l'on va de l'assignation d'une valeur de vérité au tout (ce *tout* pouvant être un ensemble de formules) à l'assignation d'une valeur de vérité aux parties.

La base de la *méthode des tableaux matriciels*, c'était les *tables de vérité* des connecteurs qui permettent d'assigner une valeur de vérité unique aux composés, une fois connue la valeur de vérité des constituants. La base de la *méthode des tableaux sémantiques* ce sont les équations de la figure 4.5 qui, *à*

*l'inverse*, permettent d'assigner une valeur *ou* une disjonction de valeurs de vérité aux constituants, une fois fixée celle des composés.

Si $\mathcal{S}$ contient l'équation $||\neg Y|| = \mathbf{V}$,
il contient aussi l'équation $||Y|| = \mathbf{F}$.

Si $\mathcal{S}$ contient l'équation $||\neg Y|| = \mathbf{F}$,
il contient aussi l'équation $||Y|| = \mathbf{V}$.

Si $\mathcal{S}$ contient l'équation $||Y \wedge Z|| = \mathbf{V}$,
il contient l'équation $||Y|| = \mathbf{V}$ et l'équation $||Z|| = \mathbf{V}$.

Si $\mathcal{S}$ contient l'équation $||Y \wedge Z|| = \mathbf{F}$,
il contient soit l'équation $||Y|| = \mathbf{F}$, soit l'équation $||Z|| = \mathbf{F}$,
soit les deux.

Si $\mathcal{S}$ contient l'équation $||Y \vee Z|| = \mathbf{V}$,
il contient soit l'équation $||Y|| = \mathbf{V}$, soit l'équation $||Z|| = \mathbf{V}$,
soit les deux.

Si $\mathcal{S}$ contient l'équation $||Y \vee Z|| = \mathbf{F}$,
il contient les équations $||Y|| = \mathbf{F}$ et $||Z|| = \mathbf{F}$.

Si $\mathcal{S}$ contient l'équation $||Y \supset Z|| = \mathbf{V}$,
il contient soit l'équation $||Y|| = \mathbf{F}$, soit l'équation $||Z|| = \mathbf{V}$,
soit les deux.

Si $\mathcal{S}$ contient l'équation $||Y \supset Z|| = \mathbf{F}$,
il contient les équations $||Y|| = \mathbf{V}$ et $||Z|| = \mathbf{F}$.

Si $\mathcal{S}$ contient l'équation $||Y \equiv Z|| = \mathbf{V}$,
il contient soit les équations $||Y|| = \mathbf{V}$ et $||Z|| = \mathbf{V}$,
soit les équations $||Y|| = \mathbf{F}$ et $||Z|| = \mathbf{F}$.

Si $\mathcal{S}$ contient l'équation $||Y \equiv Z|| = \mathbf{F}$,
il contient soit les équations $||Y|| = \mathbf{V}$ et $||Z|| = \mathbf{F}$,
soit les équations $||Y|| = \mathbf{F}$ et $||Z|| = \mathbf{V}$.

**Figure 4.5** *Equations pour les tableaux sémantiques*

La différence profonde qui sépare la méthode des tableaux sémantiques de la méthode des tables de vérité tient en ceci : la relation entre la valeur de vérité du tout et celle de ses constituants immédiats n'est pas nécessairement une *fonction* — contrairement à la relation entre la valeur de vérité des constituants

et celle du tout. Par exemple, la valeur de vérité de la formule $A \supset B$ est complètement déterminée, une fois fixée celle de $A$ et celle de $B$, mais la réciproque n'est pas vraie. Quand on sait que $A \supset B$ est vrai, on ne sait pas encore quelle est la valeur de vérité de $A$ et de $B$, car il peut y avoir plusieurs *sources* de la vérité de $A \supset B$, à savoir la vérité de $B$ ou la fausseté de $A$.

Au lieu de recourir à des règles de *composition* qui génèrent l'évaluation du tout à partir de celle des parties, comme le faisait la méthode des tables de vérité, la méthode des tableaux sémantiques a recours à des règles de *décomposition* qui fragmentent l'évaluation du tout en évaluation des parties. Certaines règles de décomposition, mais pas toutes, n'autorisent qu'*un seul résultat*. C'est le cas, par exemple, de celle qui tire les conséquences de l'équation de la figure 4.6. Nous appellerons ces règles, *règles de prolongation*

$$\begin{array}{ll} \alpha & \|p \supset q\| = \mathbf{F} \\ & \downarrow \\ \alpha.1 & \|p\| = \mathbf{V} \\ & \downarrow \\ \alpha.2 & \|q\| = \mathbf{F} \end{array}$$

**Figure 4.6** *Schéma d'une règle de prolongation*

ou règles de type $\alpha$.

Comme la *relation* de décomposition n'est pas fonctionnelle, on aura des règles qui prendront à l'entrée une évaluation de formule, c'est-à-dire une équation, et qui donneront à la sortie une *alternative* entre plusieurs évaluations de formules, c'est-à-dire plusieurs équations. Par exemple, la règle de décomposition qui prend comme *entrée* l'équation

(4.30) $\|p \supset q\| = \mathbf{V}$

aura comme *sortie* la disjonction (inclusive)

(4.31) $\|p\| = \mathbf{F}$ , $\|q\| = \mathbf{V}$,

que l'on représente graphiquement par une fourche. Nous appellerons ces règles, règles de *ramification* ou règles de type $\beta$ (la règle de la figure 4.7 se lit : "Si $p \supset q$ est vrai, alors ou $p$ est faux, ou $q$ est vrai").

Dans les présentations les plus récentes de la méthode des tableaux sémantiques (Smullyan 1968, Frost 1986), la structure d'arbre est adoptée comme dans l'algorithme d'analyse de vérité de Quine : à chaque pas, un connecteur est éliminé. Ici encore cette ressemblance de surface ne doit pas nous faire oublier la différence profonde entre les deux méthodes, différence qu'il importe de souligner ici, car elle se traduit par une réduction sensible de l'espace de recherche utilisé par la méthode des tableaux sémantiques.

Pour montrer qu'un ensemble fini de formules $\{A_1, A_2, A_3\}$ implique tau-

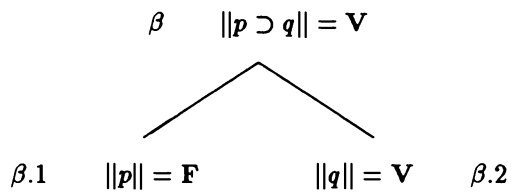



**Figure 4.7** *Schéma d'une règle de ramification*

tologiquement une formule $B$, la méthode des tableaux examine systématiquement *non pas toutes* les évaluations booléennes *indistinctement*, comme le faisait la méthode des tables de vérités, *mais seulement* toutes les évaluations booléennes qui sont susceptibles de satisfaire simultanément $A_1, A_2, A_3$ et $\neg B$. Si l'ensemble est sémantiquement inconsistant, toutes les tentatives pour trouver une évaluation booléenne qui la satisfasse échoueront. Cet échec se reconnaît à l'indice suivant : chacune des évaluations booléennes considérées contient deux équations incompatibles de la forme $\{||p|| = \mathbf{V}, ||p|| = \mathbf{F}\}$. Si l'ensemble est sémantiquement consistant, la procédure fera apparaître au moins une évaluation booléenne qui satisfera simultanément $A_1, A_2, A_3$ et $\neg B$.

La méthode des tableaux sémantiques est une méthode mécanique qui présente *deux* avantages par rapport à la méthode des tables de vérité (ou des analyses de vérité). Premièrement, elle s'applique *aussi* aux *ensembles infinis*, contrairement à celle des tables. Deuxièmement, l'*espace de recherche* est *borné* à l'ensemble des assignations de valeurs *pertinentes* pour la question posée.

### 4.1.16. *Le tableau sémantique en forme d'arbre.*

Avec Smullyan, donnons aux tableaux la forme d'un arbre. Imaginons un arbre renversé : la racine est en haut et les feuilles en bas. Il y a des nœuds à la racine, sur le tronc, sur les branches et à l'extrémité des branches. A ces nœuds sont associées des équations qui assignent une valeur à une formule, par exemple $||A \wedge B|| = \mathbf{V}$. Chaque branche représente une tentative de rendre vrais simultanément les énoncés de l'ensemble dont on teste la consistance. Dès que sur une même branche apparaissent deux évaluations mutuellement contradictoires de propositions atomiques ($||p|| = \mathbf{V}$ et $||p|| = \mathbf{F}$, que l'on abrégera en $\mathbf{V}(p)$ et $\mathbf{F}(p)$, respectivement), la tentative échoue : on dit que la branche en cause est *fermée* et que le nœud extrémité est un *nœud-échec*. On écrit une croix sous ce nœud. L'ensemble testé est consistant si et seulement si l'arbre *complet* de l'ensemble est *ouvert*. Pour définir la notion d'*arbre complet*, on a besoin des règles de décomposition.

A partir des équations booléennes énumérées plus haut, nous pouvons dresser l'inventaire de toutes les règles de décomposition de molécules, règles qui se répartissent, comme on l'a vu plus haut, en deux groupes : les règles de *prolongation* de type $\alpha$ et les règles de *ramification* de type $\beta$, dont le tableau complet se présente comme suit. Les règles de décomposition sont données à la figure 4.8.

| | | | | | | |
|---|---|---|---|---|---|---|
| ($\alpha$) | $\mathbf{V}(X \wedge Y)$ | $\mathbf{V}(\neg\neg X)$ | $\mathbf{F}(X \vee Y)$ | $\mathbf{F}(X \supset Y)$ | $\mathbf{V}(\neg X)$ | $\mathbf{F}(\neg X)$ |
| | ↓ | ↓ | ↓ | ↓ | ↓ | ↓ |
| ($\alpha$.1) | $\mathbf{V}(X)$ | $\mathbf{V}(X)$ | $\mathbf{F}(X)$ | $\mathbf{V}(X)$ | $\mathbf{F}(X)$ | $\mathbf{V}(X)$ |
| | ↓ | | ↓ | ↓ | | |
| ($\alpha$.2) | $\mathbf{V}(Y)$ | | $\mathbf{F}(Y)$ | $\mathbf{F}(Y)$ | | |

| | | | | | |
|---|---|---|---|---|---|
| ($\beta$) | $\mathbf{F}(X \wedge Y)$ | $\mathbf{V}(X \vee Y)$ | $\mathbf{V}(X \supset Y)$ | $\mathbf{V}(X \equiv Y)$ | $\mathbf{F}(X \equiv Y)$ |
| | ⟋ ⟍ | ⟋ ⟍ | ⟋ ⟍ | ⟋ ⟍ | ⟋ ⟍ |
| ($\beta_1\beta_2$) | $\mathbf{F}(X)$ $\mathbf{F}(Y)$ | $\mathbf{V}(X)$ $\mathbf{V}(Y)$ | $\mathbf{F}(X)$ $\mathbf{V}(Y)$ | $\mathbf{V}(X)$ $\mathbf{F}(X)$ | $\mathbf{F}(X)$ $\mathbf{V}(X)$ |
| | | | | ↓ ↓ | ↓ ↓ |
| | | | | $\mathbf{V}(Y)$ $\mathbf{F}(Y)$ | $\mathbf{V}(Y)$ $\mathbf{F}(Y)$ |

**Figure 4.8** *Les règles de décomposition des molécules*

Une branche $b_i$ d'un arbre est *complète* si, pour chaque prolongation $\alpha$ qui figure dans $b_i$, les deux constituants $\alpha.1$ et $\alpha.2$ se trouvent sur la branche $b_i$ et si, pour chaque ramification ($\beta$), l'un au moins des deux constituants $\beta_1$ et $\beta_2$ figurant aux extrémités de la fourche se trouve sur la branche $b_i$. Un arbre dont chaque branche est complète ou fermée est un *arbre complet*.

Dans l'exemple ci-dessous les deux branches sont complètes : pour la formule $\mathbf{V}(p \supset q)$ qui est du type $\beta$, le constituant $\beta_1$ (c'est-à-dire $\mathbf{F}p$) est sur l'une des branches et le constituant $\beta_2$ est sur l'autre. Quant à la formule $\mathbf{F}(\neg q \supset \neg p)$ qui est de type $\alpha$, ses deux constituants $\alpha_1$ (c'est-à-dire $\mathbf{V}\neg q$) et $\alpha_2$ (c'est-à-dire $\mathbf{F}\neg p$) se trouvent sur chaque branche.

On arrête de développer une branche soit s'il n'y a plus de formules non atomiques à décomposer, soit si on est arrivé à une contradiction. Si on arrive à une contradiction en écrivant la première des deux formules d'une prolongation, on achève cependant d'écrire la deuxième (c'est-à-dire $\alpha_2$) avant d'arrêter le processus.

### 4.1.17. *Construction d'arbres pour un ensemble fini*

*Initialiser* le tableau, en traçant une ligne verticale subdivisée en nœuds, et en étiquetant chaque nœud à l'aide d'une des formules de l'ensemble dont on cherche à prouver l'inconsistance. Il y a deux états terminaux : $A$ et $B$. L'arrêt en $A$ révèle que l'arbre est fini fermé et que l'ensemble des formules est

inconsistant. L'arrêt en $B$ révèle que l'arbre est fini ouvert et que l'ensemble des formules est consistant.

1. (a) Clore par une croix de fermeture chaque branche (éventuellement le tronc) qui contient deux littéraux mutuellement contradictoires. Cela doit se faire *dès* que le littéral complémentaire est apparu.
   (b) Cocher toute étiquette qui est constituée par un littéral ("$\surd$").
   (c) Si toutes les branches sont fermées, s'arrêter en $A$.
2. (a) Chercher une branche ouverte qui contient un nœud dont l'équation n'a pas encore été cochée.
   (b) Choisir un nœud $N$ sur une telle branche $L$.
   (c) Etendre toutes les branches non fermées qui passent par $N$ à l'aide d'une règle de décomposition et cocher $N$.
   (d) S'il n'y a plus de branche ouverte et incomplète, c'est-à-dire ayant au moins un nœud non coché, s'arrêter en $B$.

La procédure s'arrête[6]. Elle fait halte (H) en $A$ ou en $B$, mais pas dans les deux états.

### 4.1.18. *Exemple*

On veut prouver que

(4.32) $(p \supset q) \supset (\neg q \supset \neg p)$

est un théorème. On cherche donc à rendre fausse sa négation (voir figure 4.9).

### 4.1.19. *Une variante notationnelle de la méthode des tableaux*

La preuve par séquents a la forme d'un arbre dont la racine est en bas et les feuilles en haut. Les axiomes sont situés au niveau des feuilles, donc à l'extrémité des branches. La preuve par la méthode des tableaux a la forme d'un arbre orienté dans le sens inverse de l'arbre des séquents : la racine est en haut et le point de départ est fourni par l'ensemble des formules dont on cherche à prouver l'inconsistance.

Dans la pratique, cependant, on "construit" l'arbre d'une preuve par séquents comme celui d'une preuve par tableaux : pour prouver que l'ensemble $\beta$ est une conséquence logique de l'ensemble $\alpha$, et donc que le séquent $\alpha \Rightarrow \beta$ est valide, on cherche à vérifier l'ensemble $\alpha$ et à falsifier l'ensemble $\beta$, et c'est

[6]Démonstration : J. A. Robinson, *Logic : Form and Function*, Edinburgh, U. P. , 1979, pp. 51–52.

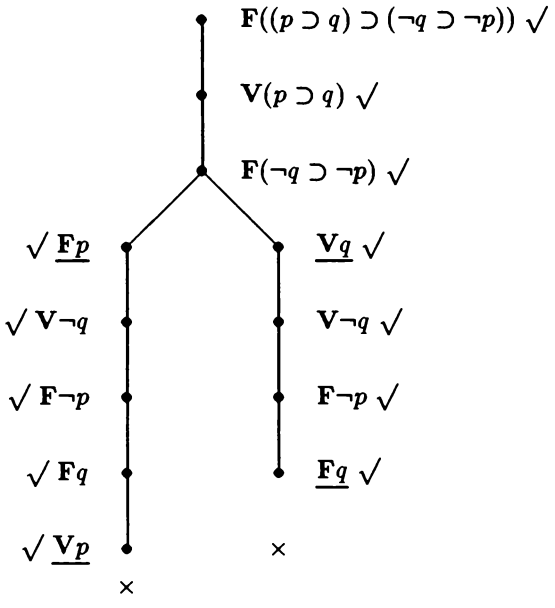



**Figure 4.9** *Tableau sémantique pour la formule 4.32*

au cours de cette quête que l'on se heurte, sur chaque branche, aux axiomes. On applique donc la maxime : "vérifier à gauche de ⇒ et falsifier à droite". La position "à gauche" est ainsi associée aux notions de vérification et de vérité et la position "de droite" à celle de falsification et de fausseté.

Massey et Frost ont eu l'idée ingénieuse de transplanter ce symbolisme dans la preuve par tableau de manière à faire l'économie des prédicats "Vrai" et "Faux". Certes, dans le tableau, il n'y a pas de flèches de séquents (⇒), mais il y a des *branches* qui partagent l'espace en zones gauche et droite; il est donc possible d'exprimer qu'une formule est à vérifier (respectivement à falsifier) sans préfixer la formule à l'aide de **V** (respectivement **F**). Il suffira de convenir que, si la formule est placée à gauche de la branche, elle est à vérifier, et que, si elle est placée à droite, elle est à falsifier.

Par exemple, la règle de prolongation $\mathbf{F}(Y \supset Z)$ ne s'écrira plus suivant la figure 4.10 mais suivant la figure 4.11.

Le même principe est adopté pour les règles de ramification.

La règle de ramification de $\mathbf{V}(Y \supset Z)$ ne s'écrira plus suivant la figure 4.12 mais suivant la figure 4.13.

Dans cette nouvelle notation, on peut obtenir un arbre beaucoup plus lisible. Reprenons l'exercice

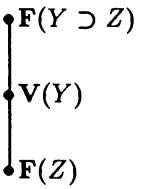



**Figure 4.10** *Prolongation classique*

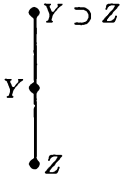



**Figure 4.11** *Prolongation nouvelle*

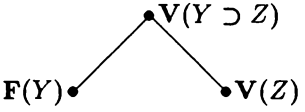



**Figure 4.12** *Ramification classique*

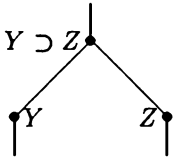



**Figure 4.13** *Ramification nouvelle*

(4.33) $s \supset (b \vee t), (b \vee a) \supset (r \wedge m), \neg r \vdash s \supset t,$

et essayons de vérifier ce qui est à gauche du $\vdash$ et de falsifier ce qui est à droite. Nous initialisons ainsi l'arbre en sachant que la position à gauche d'une branche signifie "ceci est à vérifier", et la position à droite, "ceci est à falsifier". Nous numérotons les connecteurs et les nœuds correspondants par le même chiffre. Les chiffres figurant dans les boîtes terminales nomment les branches

(cf. figure 4.14).

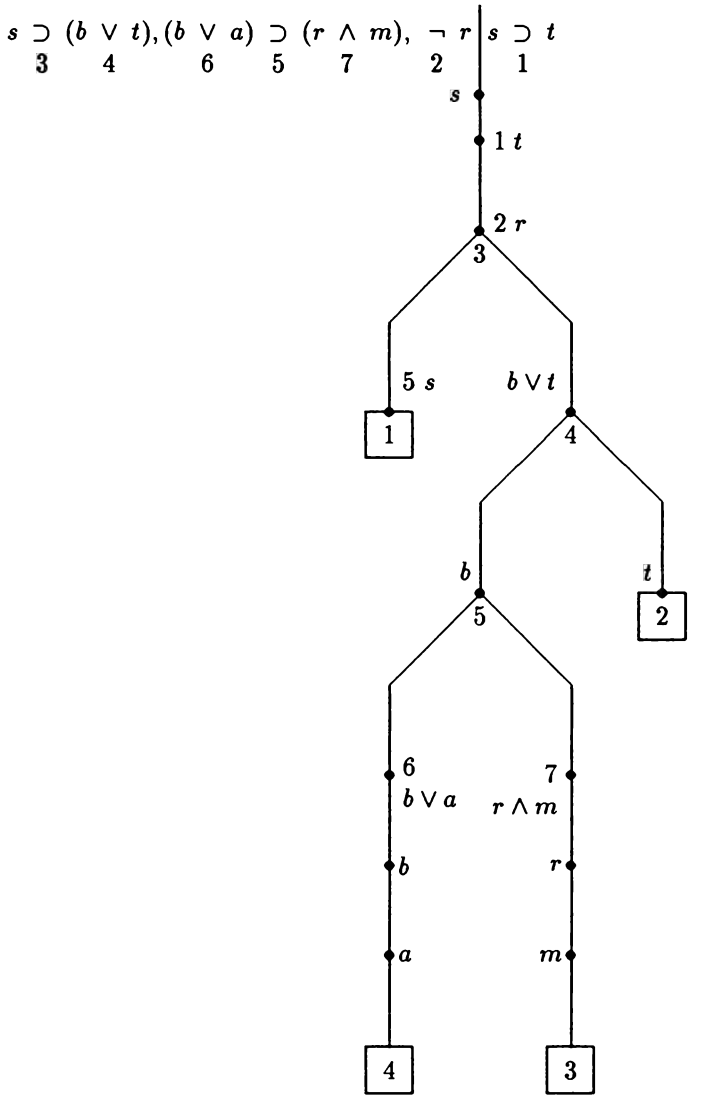



**Figure 4.14** *Tableau sémantique*

### 4.1.20. *Tableaux* versus *séquents*

On construit l'arbre de la méthode des tableaux de haut en bas, alors que l'on construit l'arbre de la méthode des séquents de bas en haut.

Dans la méthode des séquents, il y a chaque fois recopiage des formules non utilisées. On vise à remonter jusqu'à un axiome, c'est-à-dire à obtenir *sur la même ligne* deux formules atomiques d'interprétation opposée.

Au contraire, dans la méthode des tableaux, on recherche les atomes complémentaires en descendant le long d'une branche jusqu'à un noeud-échec.

Les formules apparaissant aux étapes antérieures restent disponibles. Il n'est pas nécessaire de les répéter.

La méthode des tableaux est une *méthode indirecte* : pour prouver qu'un schéma est valide, on prouve que sa négation est inconsistante. Ce genre de détour peut susciter des réticences. Comme le dit plaisamment Quine,

> la vérité est notre gibier, et une poursuite directe est chose plus franche que des lacets et des pièges (Quine, 1973, p. 205).

En revanche, elle comporte moins de redondance que la méthode des séquents.

Dans une thèse de doctorat[7] intitulée *Automated Proof Search in Non-Classical Logics : Efficient Matrix Proof Methods for Modal and Intuitionistic Logic* (1987), L. Wallen a dressé un inventaire exhaustif des imperfections que présentent la méthode des séquents et la méthode des tableaux sémantiques considérées du point de vue de l'*espace de recherche* qu'elles exigent, et il a proposé une méthode de travail plus efficiente, qui est une généralisation de la méthode des connexions de Bibel, et qui permet un traitement en machine de la logique classique, de la logique intuitionniste et des différentes logiques modales.

## 4.2. La méthode des connexions de Bibel et Wallen

Cette section peut être omise en première lecture.

### 4.2.1. *Complexité de la méthode des séquents*

1. La méthode des séquents comporte *beaucoup de redondances* : on doit à chaque pas transporter les ensembles de formules sur lesquelles on n'a pas travaillé ($\Gamma$ et $\Delta$). Cette *redondance* est éliminée en partie dans la méthode des tableaux sémantiques, mais pas complètement : même dans la méthode des tableaux, on est amené à dédoubler certaines branches.

[7] publiée en 1989 (M.I.T. Press)

2. La méthode des séquents et celle des tableaux investissent le même temps dans toutes les formules, que le traitement de celles-ci soit nécessaire pour arriver au résultat ou qu'il soit inutile. Considérons, par exemple, l'ensemble suivant dont on cherche à établir l'inconsistance :

$$\{(p \vee (q \supset r)), (p \wedge q), \neg q\}. \tag{4.34}$$

   On voit tout de suite que ce sont les deux dernières formules qui sont responsables de l'inconsistance de l'ensemble et que la première ne joue aucun rôle. L'adoption d'une routine de lecture (ici la lecture de gauche à droite) nous oblige à accorder le même temps de travail aux trois formules. On se trouve ici confronté au problème de la *pertinence*.

3. La méthode des séquents et celle des tableaux nous confrontent à des choix qui seront faits aveuglément et qui peuvent être malheureux en ce sens que la possibilité écartée nous aurait conduit *plus vite* au résultat. Nous sommes donc confrontés au problème de la *complexité*.

Comparons les deux démonstrations suivantes du syllogisme hypothétique, faites par la *méthode des séquents*.

$$\begin{array}{cccc}
 & \times & \times & \times \\
6 & q, r, p \Rightarrow r & \times & r, p \Rightarrow r, p & p \Rightarrow r, p, q \\
5 & \dfrac{}{q, r \Rightarrow (p \supset r)} & q \Rightarrow (p \supset r), q & \dfrac{}{r \Rightarrow (p \supset r), p} & \dfrac{}{\Rightarrow (p \supset r), p, q} \\
4 & \multicolumn{2}{c}{q, (q \supset r) \Rightarrow (p \supset r)} & \multicolumn{2}{c}{(q \supset r) \Rightarrow (p \supset r), p} \\
3 & \multicolumn{2}{c}{q \Rightarrow ((q \supset r) \supset (p \supset r))} & \multicolumn{2}{c}{\Rightarrow (q \supset r) \supset (p \supset r), p} \\
2 & \multicolumn{4}{c}{(p \supset q) \Rightarrow ((q \supset r) \supset (p \supset r))} \\
1 & \multicolumn{4}{c}{\Rightarrow (p \supset q) \supset ((q \supset r) \supset (p \supset r))}
\end{array}$$

**Figure 4.15** *Preuve par séquents (tactique 1)*

L'arbre de la preuve de la figure 4.15, dans laquelle on a systématiquement travaillé d'abord sur les connecteurs à gauche de la flèche $\Rightarrow$, comporte 13 nœuds (cf. figure 4.16).

Adoptons à présent la *tactique* opposée et travaillons d'abord sur les connecteurs à droite de la flèche $\Rightarrow$ (cf. figure 4.17).

L'arbre de la preuve ne comporte que 8 nœuds. La preuve est donc plus *courte*, moins *complexe*. Dans le cas présent, la deuxième tactique est plus performante que la première, mais aucun principe général n'est connu à ce jour qui nous permettrait, pour n'importe quelle formule, de savoir de quel côté du séquent il faut travailler en premier lieu pour arriver à la preuve la plus courte.

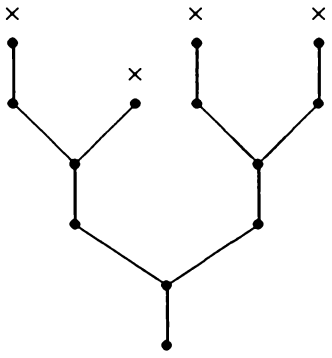


**Figure 4.16** *Arbre sans étiquettes*

$$
\begin{array}{lc}
 & \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{(p \supset q), r, p \Rightarrow r}\times \qquad \dfrac{}{p \Rightarrow r, q, p}\times \quad \dfrac{}{q, p \Rightarrow r, q}\times}{(p \supset q), p \Rightarrow r, q}}{(p \supset q), (q \supset r), p \Rightarrow r}}{(p \supset q), (q \supset r) \Rightarrow (p \supset r)}}{(p \supset q) \Rightarrow ((q \supset r) \supset (p \supset r))}}{\Rightarrow (p \supset q) \supset ((q \supset r) \supset (p \supset r))} \\
5 & \\
4 & \\
3 & \\
2 & \\
1 &
\end{array}
$$

**Figure 4.17** *Preuve par séquents (tactique 2)*

### 4.2.2. *Méthode orientée vers les connexions*

La méthode des séquents ou des tableaux — nous l'avons vu dans la section 4.1.15 — est un progrès par rapport à celle des tables de vérité en cela qu'elle consiste à examiner non pas *toutes* les évaluations booléennes indistinctement, *mais seulement* toutes les évaluations booléennes susceptibles de satisfaire la formule ou l'ensemble des formules dont on teste la consistance. Cela revient, concrètement, à se demander si toutes les branches de l'arbre généré par la méthode des séquents ou des tableaux se ferment (dans ce cas, la formule ou l'ensemble des formules est inconsistant).

On réaliserait un progrès décisif si l'on pouvait sauter l'étape du cheminement qui conduit aux atomes complémentaires, et

> remplacer la méthode qui consiste à énumérer les chemins (branches) générés par une formule et, subséquemment, à vérifier si tous les chemins tracés se ferment, par une recherche directe des connexions (atomes complémentaires) (Wallen, 1987, p. 77).

La méthode des tableaux ou des séquents est une méthode de décomposition

des formules qui est axée sur les *connecteurs*, les premiers à être traités étant ceux qui sont le plus en surface. Il est donc inévitable que *ces méthodes soient sensibles à des particularités non pertinentes de la structure de surface* (voir en 4.2.1 l'exemple de non-pertinence). L'objectif visé par Wallen est de construire une méthode orientée non pas par les *connecteurs*, mais par les *connexions* putatives que l'on cherche à trouver. Il reprend et étend la méthode des connexions inventée par BIbel (1982). *La méthode de résolution* est déjà une tentative dans ce sens, mais qui a le défaut d'exiger un travail préparatoire (la mise en forme conjonctive ou disjonctive) qui n'est pas possible dans toutes les logiques (par exemple, la logique intuitionniste). La méthode des matrices de Wallen n'exige pas ce travail préalable et est applicable aussi à des logiques non classiques comme la logique intuitionniste.

Dans le livre issu de sa thèse, Wallen résume ainsi la "méthode des matrices" :

> Le thème central est celui d'une *caractérisation par matrice*. Une telle caractérisation demande que l'on voit *géométriquement* les formules comme des structures à deux dimensions [qu'on appellera] *matrices*. A cette manière de voir est associée la notion de *chemin* parcourant la matrice et [par extension] de chemin parcourant une formule. [Ici s'introduit] l'idée de *connexion* : deux éléments (sous-formules) de la matrice (formule) sur un chemin qui sont complémentaires [qui contiennent deux littéraux complémentaires (comme $p$ et $\neg p$)] forment une "connexion". On exprime l'idée qu'une formule est un théorème en termes (de l'existence) d'un ensemble de connexions *simultanément* complémentaires qui traverse (*spans*) la formule. Cette dernière condition nous assure que chaque chemin contient une connexion puisée dans l'ensemble.
>
> Tester si une formule est un théorème se réduit à vérifier un chemin dans une matrice. Des algorithmes pour faire ce test peuvent être développés indépendamment d'une logique particulière ... (Wallen, 1989, p. 1–2)

### 4.2.3. *La codification de l'arbre syntaxique*

Dans la méthode de Wallen, on commence par générer l'*arbre syntaxique* de la formule ou de l'ensemble de formules sur la consistance desquelles on s'interroge. Cet arbre (*parsing tree*) énumère les *constituants syntaxiques* de la formule. Par exemple, si la formule examinée est le syllogisme hypothétique mentionné plus haut, on génère l'arbre de la figure 4.18.

Si l'on donne à chaque nœud un numéro de $a_0$ à $a_n$, où $a_0$ est le nœud initial, et si on lit l'arbre en appliquant le principe de numérotation en profondeur avec retour en arrière, on obtient l'arbre de la figure 4.19 qui "*codifie*" en quelque sorte le précédent. Numéroter, selon cette méthode, l'arbre de la figure 4.19 consiste à suivre l'ordre indiqué par les flèches (Delahaye, 1986, p. 170).

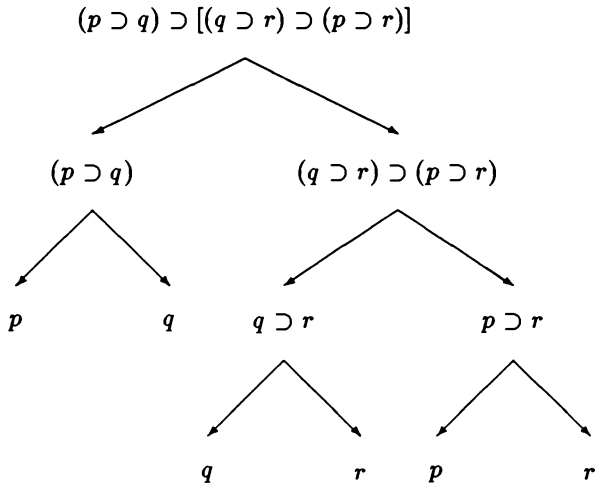



**Figure 4.18** *Arbre syntaxique*

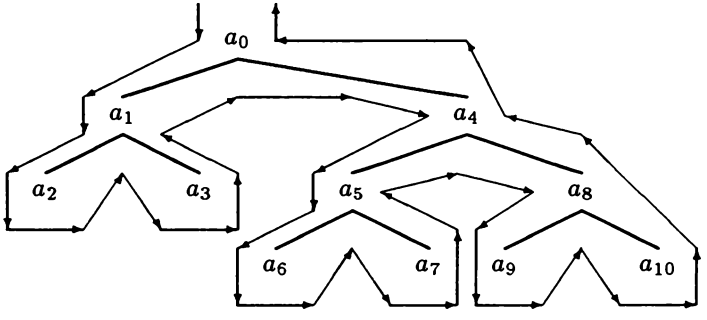



**Figure 4.19** *Arbre syntaxique numéroté*

### 4.2.4. *Définition de la polarité d'une formule*

Introduisons à présent les notions d'*occurrence positive* et d'*occurrence négative* d'une sous-formule (c'est-à-dire d'un constituant) :

> Une sous-formule $B$ [...] d'une formule donnée $A$ est dite "apparaître à titre d'occurence positive dans $A$" si elle apparaît dans la portée d'un nombre pair de négations explicites ou implicites. Dans le cas contraire, elle apparaît à titre d'occurrence négative (Wallen, 1987, p. 56).

On appelle négation explicite, le symbole de la négation $\neg$. Avoir une négation implicite, c'est être l'antécédent d'un conditionnel (en effet, $p \supset q$ équivaut à

$\neg p \vee q$), ou être la phrase à falsifier (racine de l'arbre). Le constituant $q \supset r$ dans le séquent $p \supset q \Rightarrow (q \supset r) \supset (p \supset r)$ apparaît à titre d'occurrence positive : comme antécédent d'un conditionnel, il apparaît négativement, mais comme il est à droite d'une flèche de séquent, son signe change à nouveau. En effet, on se souvient qu'on met à droite de $\Rightarrow$ la disjonction (ici réduite à un membre) des $B_1, \ldots, B_k$ à falsifier. On appelle *polarité* le signe des constituants. En parcourant l'arbre syntaxique, on voit tout de suite quelle polarité (valeur de vérité) on doit attribuer à chaque constituant pour falsifier la formule que l'on teste. Cette assignation de polarité est la deuxième opération à accomplir.

### 4.2.5. *Règles de prolongation et de ramification*

Une fois que l'on connaît le connecteur principal d'un constituant ainsi que sa polarité, on sait quelle règle d'élimination des connecteurs on doit appliquer. Ce sont les règles usuelles du calcul des séquents.

Appelons règles de type $\alpha$ les règles de *prolongation* : leurs descendants seront de type $\alpha_1$ et $\alpha_2$. Appelons règles de type $\beta$ les règles de *ramification* : leurs descendants seront de type $\beta_1$ et $\beta_2$.

| $\alpha$ | $\alpha_1$ | $\alpha_2$ |
|---|---|---|
| $\langle X \wedge Y, \mathbf{V}\rangle$ | $\langle X, \mathbf{V}\rangle$ | $\langle Y, \mathbf{V}\rangle$ |
| $\langle X \vee Y, \mathbf{F}\rangle$ | $\langle X, \mathbf{F}\rangle$ | $\langle Y, \mathbf{F}\rangle$ |
| $\langle X \supset Y, \mathbf{F}\rangle$ | $\langle X, \mathbf{V}\rangle$ | $\langle Y, \mathbf{F}\rangle$ |
| $\langle \neg X, \mathbf{V}\rangle$ | $\langle X, \mathbf{F}\rangle$ | $\langle X, \mathbf{F}\rangle$ |
| $\langle \neg X, \mathbf{F}\rangle$ | $\langle X, \mathbf{V}\rangle$ | $\langle X, \mathbf{V}\rangle$ |

| $\beta$ | $\beta_1$ | $\beta_2$ |
|---|---|---|
| $\langle X \wedge Y, \mathbf{F}\rangle$ | $\langle X, \mathbf{F}\rangle$ | $\langle Y, \mathbf{F}\rangle$ |
| $\langle X \vee Y, \mathbf{V}\rangle$ | $\langle X, \mathbf{V}\rangle$ | $\langle Y, \mathbf{V}\rangle$ |
| $\langle X \supset Y, \mathbf{V}\rangle$ | $\langle X, \mathbf{F}\rangle$ | $\langle Y, \mathbf{V}\rangle$ |

**Figure 4.20** *Inventaire des règles*

### 4.2.6. *Types de formules*

Comme dans la méthode des tableaux et dans la méthode des séquents, on distingue entre les formules qui donnent naissance à une *prolongation*, de celles qui donnent naissance à une *ramification*. Chaque formule appartient à deux types (sauf les formules atomiques et la formule initiale) : un type primaire et un type secondaire : la formule $X$ a le type *primaire* $\alpha$ si elle donne

naissance à une *prolongation*, et le type *primaire* $\beta$ si elle donne naissance à une ramification. Une formule $X$ a le type *secondaire* $\alpha_1$ ou $\alpha_2$, si elle descend de la formule-mère par la prolongation, et le type *secondaire* $\beta_1$ ou $\beta_2$, si elle en descend par ramification.

### 4.2.7. *Arbre de formule indexé*

Après avoir inventorié tous les constituants d'une formule grâce à l'arbre syntaxique, on spécifie la polarité et les types primaires et secondaires des constituants dans un arbre de formule indexé. Par exemple, pour la formule à falsifier

(4.35) $(p \supset q) \supset [(q \supset r) \supset (p \supset r)]$,

le tableau récapitulatif se présente selon la figure 4.21.

| $k$ | pol($k$) | étiquette($k$) | type prim. | type sec. |
|---|---|---|---|---|
| $a_0$ | **F** | $(p \supset q) \supset [(q \supset r) \supset (p \supset r)]$ | $\alpha$ | – |
| $a_1$ | **V** | $p \supset q$ | $\beta$ | $\alpha_1$ |
| $a_2$ | **F** | $p$ | – | $\beta_1$ |
| $a_3$ | **V** | $q$ | – | $\beta_2$ |
| $a_4$ | **F** | $(q \supset r) \supset (p \supset r)$ | $\alpha$ | $\alpha_2$ |
| $a_5$ | **V** | $q \supset r$ | $\beta$ | $\alpha_1$ |
| $a_6$ | **F** | $q$ | – | $\beta_1$ |
| $a_7$ | **V** | $r$ | – | $\beta_2$ |
| $a_8$ | **F** | $p \supset r$ | $\alpha$ | $\alpha_2$ |
| $a_9$ | **V** | $p$ | – | $\alpha_1$ |
| $a_{10}$ | **F** | $r$ | – | $\alpha_2$ |

**Figure 4.21** *Arbre de formule indexé*

### 4.2.8. *La notion de chemin*

Wallen introduit ensuite la notion de chemin au moyen de la définition récursive suivante :

**Base :** si $k_0$ est la racine, $k_0$ est un chemin.

**Récursion :** si $S$ est un chemin contenant le noeud $\alpha^k$, $(S\backslash\{\alpha^k\}), \alpha_1^k, \alpha_2^k$ est un chemin. En d'autres termes, si on est face à un nœud $\alpha^k$ de type $\alpha$ (prolongation), on continue le chemin en soustrayant $\alpha^k$ et en ajoutant ses deux descendants, qui sont respectivement de type secondaire $\alpha_1$ et $\alpha_2$.

Si $S$, $\beta^k$ est un chemin, $(S\backslash\{\beta^k\}), \beta_1^k$ et $(S\backslash\{\beta^k\}), \beta_2^k$ sont des chemins.

En d'autres termes, si on est face à un nœud $\beta^k$ de type $\beta$ (ramification), on soustrait $\beta^k$ et on crée une bifurcation, en ajoutant à chacune des branches de la fourche l'un des descendants qui est de type secondaire $\beta_1$ ou $\beta_2$.

En appliquant cette définition récursive à la formule occupant la position $a_0$ dans la matrice, on génère un arbre ***unique*** dont les extrémités sont des ***suites d'atomes***. C'est la troisième et dernière tâche à accomplir. Ces suites peuvent coïncider en tout ou en partie. Si ***toutes*** ces suites d'atomes contiennent une connexion, c'est-à-dire deux atomes complémentaires (atomes de même lettre, mais de signe opposé), alors la formule initiale (ou l'ensemble de formules) est inconsistante. Dans l'exemple considéré, les chemins sont décrits à la figure 4.22. Quand on a tracé les chemins, on inspecte chaque extrémité pour voir s'il s'y trouve des atomes complémentaires. Si c'est le cas pour chacune de celles-ci, l'inconsistance est démontrée.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1) | | | | $\{a_0\}$ | | | |
| 2) | | | | $\{a_1, a_4\}$ | | | |
| 3) | | $\{a_2, a_4\}$ | | | | $\{a_3, a_4\}$ | |
| 4) | | $\{a_2, a_5, a_8\}$ | | | | $\{a_3, a_5, a_8\}$ | |
| 5) | $\{a_2, a_6, a_8\}$ | | $\{a_2, a_7, a_8\}$ | | $\{a_3, a_6, a_8\}$ | | $\{a_3, a_7, a_8\}$ |
| 6) | $\{a_2, a_6, a_9, a_{10}\}$ | | $\{a_2, a_7, a_9, a_{10}\}$ | | $\{a_3, a_6, a_9, a_{10}\}$ | | $\{a_3, a_7, a_9, a_{10}\}$ |

Chacune des extrémités des quatre chemins de la formule contient au moins une connexion :
$\{\mathbf{F}_p, \mathbf{V}_p \ldots\}$ $\{\mathbf{F}_p, \mathbf{V}_p, \mathbf{V}_r, \mathbf{F}_r\}$ $\{\mathbf{V}_q, \mathbf{F}_q \ldots\}$ $\{\mathbf{V}_r, \mathbf{F}_r \ldots\}$
ou, en termes de positions : $\{a_2, a_9\}$ $\{a_2, a_9, a_7, a_{10}\}$ $\{a_3, a_6\}$ $\{a_7, a_{10}\}$

**Figure 4.22** *Chemins et connexion*

### 4.2.9. *Evaluation de la méthode des connexions*

Les trois défauts de la méthode des séquents et de celle des tableaux ont été éliminés : a) il n'y a plus de *redondance* : aucune formule n'est répétée, on répète uniquement des pointeurs $a_k$ vers les formules ($a_k$ n'est pas une formule mais une *position*), b) il n'y a plus de recherche *non pertinente* : on s'occupe uniquement des connexions et on ignore les cheminements variés par lesquels on aboutit à des atomes, c) la *complexité* est réduite. On n'a plus à choisir entre *plusieurs* arbres de preuve possibles : on travaille sur l'arbre syntaxique, qui est *unique*, contrairement aux arbres de preuve par séquents, ou aux tableaux. Il s'agit là d'un *progrès majeur*. L'indéterminisme lié à la pluralité des arbres de preuve et des tableaux est éliminé.

### 4.2.10. *Application de la méthode des connexions*

Reprenons à nouveau le raisonnement

(4.36) $s \supset (b \vee t), (b \vee a) \supset (r \wedge m), \neg r \vdash (s \supset t)$.

Nous présenterons successivement

- l'arbre d'analyse (cf. figures 4.23 et 4.24),
- la classification au point de vue de la polarité et des types primaires et secondaires (cf. figure 4.25),
- les extrémités des chemins et les connexions (cf. figure 4.26).

Nous épargnerons au lecteur la construction de l'arbre des chemins qui n'est utile que pour un traitement sur ordinateur.

## 4.3. La méthode de réduction à la forme normale

Rappelons qu'on appelle *littéral* un atome ($q$) ou sa négation ($\overline{q}$). Les littéraux $q$ et $\overline{q}$ sont appelés *complémentaires* (l'un de l'autre).

Une disjonction de conjonctions de littéraux ($\ell_i$) est appelée *forme disjonctive normale* (FDN) :

(4.37) $(\ell_1 \wedge \cdots \wedge \ell_n) \vee \cdots \vee (\ell_1 \wedge \cdots \wedge \ell_m)$.

Une conjonction de disjonctions de littéraux est appelée *forme conjonctive normale* (FCN) :

(4.38) $(\ell_1 \vee \cdots \vee \ell_n) \wedge \cdots \wedge (\ell_1 \vee \cdots \vee \ell_m)$.

Des expressions de la forme :

$$\ell_1 \vee \cdots \vee \ell_i,$$

$$\ell_1 \wedge \cdots \wedge \ell_j,$$

sont des cas dégénérés que l'on peut ranger aussi bien dans la rubrique des FDN que dans celle des FCN.

Une expression est en *forme disjonctive normale développée* (FDND, en anglais *extended disjunctive normal form*) si elle est en FDN et si chaque énoncé atomique dans la forme d'origine figure une et une seule fois dans chaque membre de la *disjonction*. La *forme conjonctive normale développée* (FCND) se définit de la même manière, *mutatis mutandis*.

Pour "développer" une formule, on utilise, s'il s'agit d'une forme disjonctive normale, la loi $p \equiv (p \wedge (q \vee \overline{q})) \equiv ((p \wedge q) \vee (p \wedge \overline{q}))$, et s'il s'agit d'une forme conjonctive normale, la loi $p \equiv (p \vee (q \wedge \overline{q})) \equiv ((p \vee q) \wedge (p \vee \overline{q}))$. Ensuite, on restaure l'ordre alphabétique s'il s'est perdu, et on élimine les répétitions de constituants s'il y en a.

Toute formule du calcul des propositions peut être réduite mécaniquement à chacune de ces quatre formes. Au cours de la réduction, on ne peut utiliser

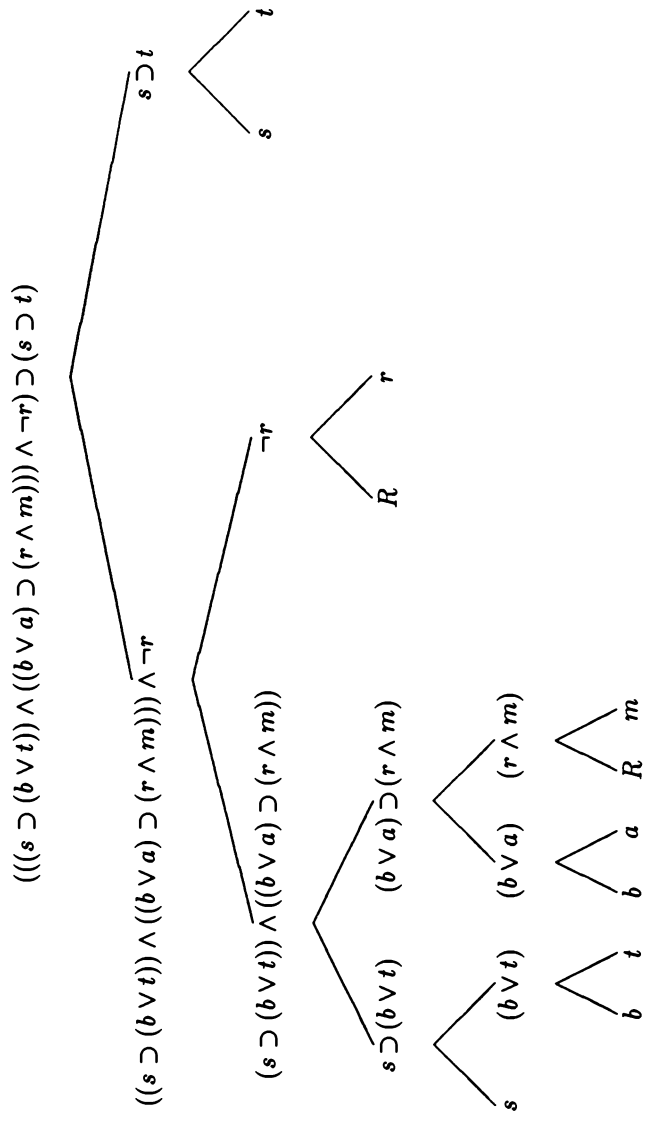



**Figure 4.23** *Arbre d'analyse de la formule 4.36*

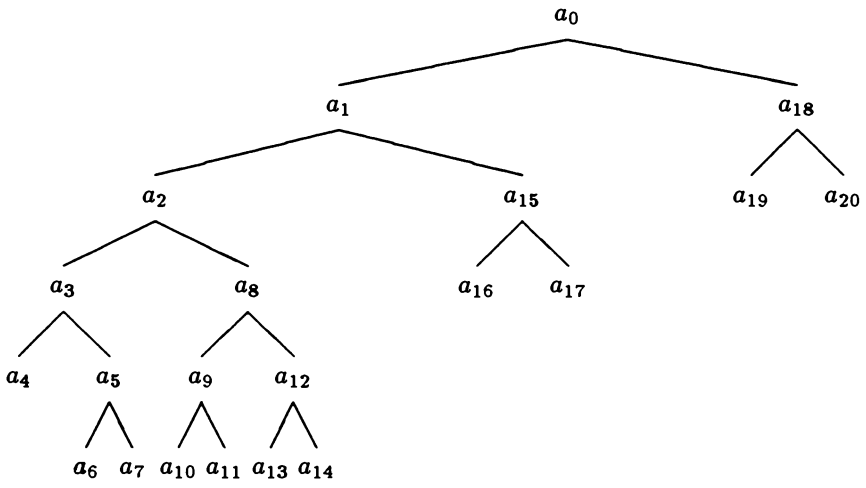



**Figure 4.24** *Arbre étiqueté*

que des transformations qui préservent l'équivalence, c'est-à-dire des transformations réversibles (par exemple la loi de De Morgan ou les principes de distributivité, mais *pas* le *modus ponens*). Pour mettre en forme disjonctive (respectivement conjonctive) normale, on procède comme suit.

1. On élimine les symboles $\equiv$ et $\supset$ grâce aux lois suivantes :

$$\begin{array}{lll} p \equiv q & \text{est logiquement équivalent à} & (p \supset q) \wedge (q \supset p) \\ p \supset q & \text{"} & (\neg p \vee q). \end{array}$$

2. On réduit la portée de la négation grâce aux lois de De Morgan :

$$\begin{array}{lll} \neg(p \vee q) & \text{est logiquement équivalent à} & (\neg p \wedge \neg q) \\ \neg(p \wedge q) & \text{"} & (\neg p \vee \neg q). \end{array}$$

3. On distribue la conjonction sur la disjonction (respectivement la conjonction sur la disjonction) grâce aux principes de distribution, où nous écrivons $pq$ pour $(p \wedge q)$ :

$$\begin{array}{lll} p \wedge (q \vee r) & \text{est logiquement équivalent à} & pq \vee pr \\ (p \vee q) \wedge r & \text{"} & pr \vee qr \\ p \vee qr & \text{"} & (p \vee q) \wedge (p \vee r) \\ pq \vee r & \text{"} & (p \vee r) \wedge (q \vee r). \end{array}$$

L'intérêt de cette réduction réside dans le fait que l'on peut établir par inspection l'inconsistance d'une expression en FDN, la validité d'une forme

| $k$ | pol($k$) | lab($k$) | Ptype($k$) | Stype($k$) |
|---|---|---|---|---|
| $a_0$ | **F** | $(((s \supset (b \vee t)) \wedge ((b \vee a) \supset (r \wedge m))) \wedge \neg r) \supset (s \supset t)$ | $\alpha$ | – |
| $a_1$ | **V** | $(((s \supset (b \vee t)) \wedge ((b \vee a) \supset (r \wedge m))) \wedge \neg r)$ | $\alpha$ | $\alpha_1$ |
| $a_2$ | **V** | $(s \supset (b \vee t)) \wedge ((b \vee a) \supset (r \wedge m))$ | $\alpha$ | $\alpha_1$ |
| $a_3$ | **V** | $s \supset (b \vee t)$ | $\beta$ | $\alpha_1$ |
| $a_4$ | **F** | $s$ | – | $\beta_1$ |
| $a_5$ | **V** | $b \vee t$ | $\beta$ | $\beta_2$ |
| $a_6$ | **V** | $b$ | – | $\beta_1$ |
| $a_7$ | **V** | $t$ | – | $\beta_2$ |
| $a_8$ | **V** | $(b \vee a) \supset (r \wedge m)$ | $\beta$ | $\alpha_2$ |
| $a_9$ | **F** | $(b \vee a)$ | $\alpha$ | $\beta_1$ |
| $a_{10}$ | **F** | $b$ | – | $\alpha_1$ |
| $a_{11}$ | **F** | $a$ | – | $\alpha_2$ |
| $a_{12}$ | **V** | $r \wedge m$ | $\alpha$ | $\beta_2$ |
| $a_{13}$ | **V** | $r$ | – | $\alpha_1$ |
| $a_{14}$ | **V** | $m$ | – | $\alpha_2$ |
| $a_{15}$ | **V** | $\neg r$ | $\alpha$ | $\alpha_2$ |
| $a_{16}$ | **F** | $r$ | – | $\alpha_1$ |
| $a_{17}$ | **F** | $r$ | – | $\alpha_2$ |
| $a_{18}$ | **F** | $s \supset t$ | $\alpha$ | $\alpha_2$ |
| $a_{19}$ | **V** | $s$ | – | $\alpha_1$ |
| $a_{20}$ | **F** | $t$ | – | $\alpha_2$ |

**Figure 4.25** *Arbre indexé de formule*

| | | |
|---|---|---|
| 1) | $a_4, a_{10}, a_{11}, a_{16}, a_{17}, a_{19}, a_{20}$ | $a_4 - a_{19} : s^{\mathbf{F}}, s^{\mathbf{V}}$ |
| 2) | $a_6, a_{10}, a_{11}, a_{16}, a_{17}, a_{19}, a_{20}$ | $a_6 - a_{10} : b^{\mathbf{V}}, b^{\mathbf{F}}$ |
| 3) | $a_7, a_{10}, a_{11}, a_{16}, a_{17}, a_{19}, a_{20}$ | $a_7 - a_{20} : t^{\mathbf{V}}, t^{\mathbf{F}}$ |
| 4) | $a_4, a_{13}, a_{14}, a_{16}, a_{17}, a_{19}, a_{20}$ | $a_4 - a_{19} : s^{\mathbf{F}}, s^{\mathbf{V}}$ |
| 5) | $a_6, a_{13}, a_{14}, a_{16}, a_{17}, a_{19}, a_{20}$ | $a_{13} - a_{16} : r^{\mathbf{V}}, r^{\mathbf{F}}$ |
| 6) | $a_7, a_{13}, a_{14}, a_{16}, a_{17}, a_{19}, a_{20}$ | $a_7 - a_{20} : t^{\mathbf{V}}, t^{\mathbf{F}}$ |

**Figure 4.26** *Extrémités des chemins et connexions*

en FCN, ainsi que l'inconsistance, la simple consistance et la validité d'une expression en FDND ou en FCND, mais le coût de la réduction elle-même est élevé.

Le schéma $S$ en FDN est inconsistant si et seulement si après simplification graduelle des schémas visiblement inconsistants, on se trouve, pour terminer, en présence d'une disjonction d'expressions manifestement inconsistante, c'est-à-dire d'une disjonction de conjonctions qui, *toutes*, contiennent deux littéraux complémentaires.

$$(p \supset q) \wedge (p\overline{q})$$
$$(\overline{p} \vee q) \wedge (p\overline{q})$$
$$\overline{p}p\overline{q} \vee qp\overline{q}$$

**Figure 4.27** *Réduction en forme disjonctive normale : formule inconsistante*

Le schéma $S$ en FCN est tautologique si et seulement si après les simplifications graduelles de constituants tautologiques, on se trouve face à une conjonction manifestement tautologique, c'est-à-dire une conjonction de disjonctions qui, *toutes*, contiennent deux littéraux complémentaires.

$$\neg((p \supset q) \wedge (p\overline{q}))$$
$$\neg(p \supset q) \vee \neg(p\overline{q})$$
$$(p\overline{q}) \vee (\overline{p} \vee q)$$
$$(p\overline{q} \vee \overline{p}) \vee q$$
$$((\overline{p} \vee p) \wedge (\overline{p} \vee \overline{q})) \vee q$$
$$(\overline{p} \vee p \vee q) \wedge (\overline{p} \vee \overline{q} \vee q)$$

**Figure 4.28** *Réduction en forme disjonctive normale : formule valide*

Un schéma $S$ en forme disjonctive normale développée est

- *tautologique* si, pour $n$ lettres, il a $2^n$ constituants,
- *inconsistant* si, pour $n$ lettres, il a 0 constituant,
- *simplement consistant* si, pour $n$ lettres, il a $x$ constituants où $0 < x < 2^n$.

*Exemple de mise en FDND*

Soit le schéma inconsistant en FDN qui contient deux variables propositionnelles distinctes $p$ et $q$ :

(4.39) $\quad \overline{p}q\overline{q} \vee qp\overline{q}.$

Pour qu'il soit en FDND, il faudrait qu'*aucun* des membres de la disjonction ne contienne ni *plus* d'une occurrence de $p$ et d'une occurrence de $q$ ni *moins* d'une occurrence de $p$ et d'une occurrence de $q$. A la limite, cette exigence est satisfaite si le schéma ne contient aucun membre. En effet, s'il n'en contient aucun, il n'en contient aucun qui déroge à la règle. C'est ce qui se passe ici. En vertu de la règle de réécriture de l'algorithme de Quine ($A \vee \mathbf{F} \rightarrow A$), tous les constituants inconsistants disparaissent. Les deux constituants $\overline{p}q\overline{q}$ et $qp\overline{q}$ qui dérogeaient disparaissent.

Soit le schéma valide

(4.40) $\quad \neg((p \supset q) \wedge (p\overline{q})).$

Mettons-le en FDN :

(4.41) $\quad \neg(\overline{p} \vee q) \vee \neg(p\overline{q})$
$\qquad p\overline{q} \vee \overline{p} \vee q,$

puis en FDND :

(4.42) $\quad p\overline{q} \vee \overline{p}(q \vee \overline{q}) \vee q(p \vee \overline{p})$
$\qquad p\overline{q} \vee \overline{p}q \vee \overline{p}\,\overline{q} \vee pq$
$\qquad pq \vee p\overline{q} \vee \overline{p}q \vee \overline{p}\,\overline{q}.$

Le schéma comporte bien $2^2$ constituants.

Un schéma $S$ en forme conjonctive normale développée est

- *tautologique* si, pour $n$ lettres, il a 0 constituant,
- *inconsistant* si, pour $n$ lettres, il a $2^n$ constituants,
- *simplement consistant* si, pour $n$ lettres, il a $x$ constituants où $0 < x < 2^n$.

Soit le schéma valide en FCN qui contient deux variables propositionnelles distinctes. Mis en FCND, il comportera 0 constituant. En effet, en vertu de la règle de réécriture de l'algorithme de Quine ($A \wedge \mathbf{V} \rightarrow A$), tous les constituants tautologiques disparaissent.

Soit le schéma inconsistant

(4.43) $\quad (p \supset q) \wedge (p\overline{q}).$

Mettons-le en FCN :

(4.44) $(\overline{p} \vee q) \wedge p \wedge \overline{q}$,

puis en FCND :

(4.45) $$(\overline{p} \vee q) \wedge p \vee (q\overline{q}) \wedge \overline{q} \vee (p\overline{p})$$
$$(\overline{p} \vee q) \wedge (p \vee q) \wedge (p \vee \overline{q}) \wedge (\overline{p} \vee \overline{q})$$
$$(p \vee q) \wedge (p \vee \overline{q}) \wedge (\overline{p} \vee q) \wedge (\overline{p} \vee \overline{q}).$$

Le schéma comporte bien $2^2$ constituants.

## 4.4. La méthode de résolution

### 4.4.1. *Algorithmes pour l'homme et pour la machine*

La méthode des tableaux est une procédure mécanique n'exigeant pas d'imagination. Elle n'est cependant pas le meilleur algorithme pour la machine. Les patrons d'inférence les plus adaptés à l'homme font appel à un grand nombre de petites inférences simples et dont la portée est claire. Pour la machine, mieux vaut un patron d'inférence comportant une seule règle, fût-elle complexe. Ce qui est difficile à traiter pour un homme peut être facile pour la machine et vice-versa.

Beaucoup de démonstrateurs automatiques de théorèmes utilisent comme unique règle d'inférence la *règle de résolution*.

### 4.4.2. *La notion de clause*

On appelle *clause* une formule constituée d'une disjonction de littéraux. Mettre une formule en *forme clausale*, c'est tout simplement la mettre en forme conjonctive normale.

### 4.4.3. *La règle de résolution*

La règle prend comme *entrée* une *paire* de clauses $C_1$ et $C_2$ telle que $C_1$ et $C_2$ contiennent respectivement les littéraux complémentaires $\ell$ et $\overline{\ell}$, et comme *sortie* la clause formée par les membres restants des deux clauses. Nous représenterons par $C_i'$ la disjonction des membres restant de $C_i$. On appelle *résolvant* la clause issue de $C_1$ et $C_2$ de cette manière.

Formellement la règle s'énonce ainsi, où $l$ désigne un littéral positif $(p, q, r, \ldots)$ et $\overline{\ell}$ un littéral négatif $(\overline{p}, \overline{q}, \overline{r}, \ldots)$ :

(4.46) $$\underbrace{(C_1' \vee \ell)}_{C_1} \wedge \underbrace{(C_2' \vee \overline{\ell})}_{C_2} \vdash \underbrace{C_1' \vee C_2'}_{R}.$$

*Exemple.*

(4.47) $(\overline{s} \vee b \vee t) \wedge (\overline{b} \vee r) \vdash (\overline{s} \vee t \vee r)$.

**Théorème 4.1 : correction de la règle de résolution.** La relation syntaxique de dérivabilité recouvre une relation sémantique de conséquence logique. La règle est donc correcte.

*Démonstration.* Supposons que $||C_1||_{\mathcal{I}} = \mathbf{V}$ et $||C_2||_{\mathcal{I}} = \mathbf{V}$. Dans cette même interprétation, $||\ell|| = \mathbf{F}$ ou $||\overline{\ell}|| = \mathbf{F}$. Si $||\ell|| = \mathbf{F}$, alors $C_1$ ne peut être vrai que si $C_1'$ est vrai. Dans ce cas, on a par addition $(p \supset p \vee q), ||C_1' \vee C_2'|| = \mathbf{V}$. Si $||\overline{\ell}|| = \mathbf{F}$, alors par le même raisonnement $||C_2'|| = \mathbf{V}$ et donc aussi $||C_1' \vee C_2'|| = \mathbf{V}$.

Si l'on ajoute à un ensemble de clauses $E$ la clause résolvante $R$ , le nouvel ensemble $E'$ obtenu de cette manière est *logiquement équivalent* à l'ensemble $E$. Cette opération d'adjonction est itérable. Elle se fonde sur le fait que la formule 4.48 est valide (Gallier, 1986, p. 128).

(4.48) $\models ((C_1' \vee \ell) \wedge (C_2' \vee \overline{\ell})) \equiv ((C_1' \vee \ell) \wedge (C_2' \vee \overline{\ell}) \wedge (C_1' \vee C_2'))$

Si $E'$ est logiquement équivalent à $E$, $E'$ est *valide* (respectivement *simplement consistant*, *inconsistant*) si et seulement si $E$ est *valide* (respectivement *simplement consistant*, *inconsistant*).

### 4.4.4. *La procédure de résolution par saturation de niveau*

De là découle une méthode mécanique simple pour prouver qu'un ensemble de clauses est inconsistant :

> Une manière directe d'effectuer la résolution sur l'ensemble $E$, est de calculer tous les résolvants de paires de clauses dans $E$, d'ajouter ces résolvants à l'ensemble $E$, d'effectuer la résolution sur le nouvel ensemble et de répéter le processus jusqu'à ce que nous trouvions une paire de littéraux complémentaires que nous noterons $\mathcal{F}$ (le symbole $\mathcal{F}$ désigne la valeur de vérité **F**). En d'autres termes, nous générons les séquences $E^0, E^1, E^2 \ldots$ où
>
> $$\begin{aligned} E^0 &= E \\ E^n &= \{\text{résolvants de } C_1 \text{ et } C_2 : C_1 \in (E^0 \cup \ldots \cup E^{n-1}) \\ &\qquad \text{et } C_2 \in E^{n-1}\},\ n = 1, 2, 3 \ldots \end{aligned}$$
>
> (Chang et Lee, 1973, p. 92).

*Exemple.* Reprenons une nouvelle fois la preuve de

(4.49) $\{s \supset (b \vee t), (b \vee a) \supset (r \wedge m), \overline{r}\} \vdash s \supset t.$

Pour prouver que $s \supset t$ découle bien de l'ensemble $E = \{s \supset (b \vee t), (b \vee a) \supset (r \wedge m), \overline{r}\}$, nous formons l'ensemble $E' = E \cup \{\neg(s \supset t)\}$ et nous cherchons à établir

(4.50) $E' \vdash \mathcal{F}.$

On commence par mettre les formules membres de $E'$ en forme clausale. On obtient

$$E' = \{\underbrace{(\overline{s} \vee b \vee t)}_{C_1}, \underbrace{(\overline{b} \vee r)}_{C_2}, \underbrace{(\overline{b} \vee m)}_{C_3}, \underbrace{(\overline{a} \vee r)}_{C_4}, \underbrace{(\overline{a} \vee m)}_{C_5}, \underbrace{\overline{r}}_{C_6}, \underbrace{s}_{C_7}, \underbrace{\overline{t}}_{C_8}\}. \tag{4.51}$$

On résout les clauses $C_1$ et $C_2$, ce qui donne le résolvant

$$(\overline{s} \vee t \vee r), \tag{4.52}$$

que l'on ajoute à $E'$ de manière à obtenir l'ensemble $E''$. On continue à résoudre et à ajouter le résolvant à l'ensemble précédent, jusqu'au moment où on arrive à la clause vide (cf. figure 4.29).

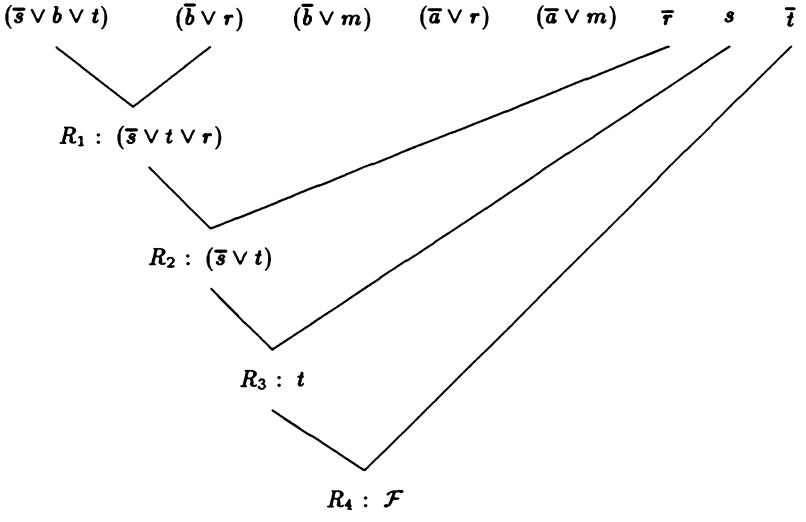



**Figure 4.29** *Preuve de 4.49 par résolution*

### 4.4.5. *Coût de la méthode en espace de recherche*

Si l'on applique la méthode de résolution par saturation à l'ensemble de clauses $E^0 = \{p \vee q, \neg p \vee q, p \vee \neg q, \neg p \vee \neg q\}$ de manière mécanique, on n'arrive à générer la clause vide qu'après avoir exécuté 35 résolutions (Chang et Lee, pp. 92–93). Cette lenteur provient du fait que l'on génère chemin faisant des *tautologies*, qui ne contribuent en rien à rendre l'ensemble inconsistant, ainsi que des clauses *redondantes*. On améliore considérablement les performances de la méthode, si on adopte la stratégie qui consiste à *effacer* les tautologies ainsi que les clauses qui sont sous d'autres clauses chaque fois qu'on en rencontre. Les tautologies se voient du premier regard puisqu'on travaille avec des expressions sous forme clausale. Quant aux clauses redondantes, c'est-à-dire subsumées

sous d'autres clauses, on les détecte à l'aide d'un algorithme (Chang et Lee, *Ibid.*, pp. 95–96).

# Chapitre 5
# Le calcul des prédicats
# – méthodes synthétiques –

## 5.1. Introduction

Dans *Logique mathématique* (1967, 1972), Kleene s'assignait pour tâche de donner un traitement complet quoique élémentaire de la logique du premier ordre et jugeait nécessaire de ne pas se limiter à une seule formulation.

> Il est caractéristique des logiciens contemporains de travailler avec souplesse, en utilisant plusieurs formulations et en passant de l'une à l'autre selon les exigences de leur propos (Kleene, 1972, 7).

C'est l'optique que nous avons adoptée, dans les chapitres 3 et 4, en présentant et en comparant huit méthodes de preuve. Deux d'entre elles cessent d'être applicables : il s'agit de la méthode des tableaux matriciels et de la méthode de la réduction à la forme normale. Les six autres méthodes peuvent être étendues au calcul des prédicats, à condition que l'on enrichisse les règles de formation, les axiomes, les règles d'inférence. Il faut aussi renforcer les contraintes imposées aux arbres de preuve. Nous nous concentrerons dans le chapitre 5 sur ce qu'il faut *ajouter* au calcul des propositions pour obtenir un calcul des prédicats.

## 5.2. Syntaxe du calcul des prédicats

Le calcul des prédicats est une extension du calcul des propositions :

> [La logique des prédicats] contient en elle-même toute l'algèbre des propositions, c'est-à-dire les propositions [...] pouvant prendre les valeurs V et

> **F**, toutes les opérations de l'algèbre des propositions et, par conséquent, toutes ses formules. Mais en outre, la logique des prédicats introduit, dans les propositions considérées, des attributs aux objets. Il y a déjà, en elle, une division des propositions en sujet et prédicat. (Novikov, 1964, p.99)

Il importe de distinguer soigneusement l'opposition entre *termes singuliers* et *termes généraux* introduite en 1.4.1, de l'opposition entre *sujet* et *prédicat*. En effet, la juxtaposition d'un terme général tel que "impair" et d'un terme singulier tel que "5" ne suffit pas à former une proposition. Elle ne constitue encore qu'une liste (cf 1.4.4). Au contraire la juxtaposition du terme singulier "5" pris comme sujet et du prédicat "est impair" forme une proposition. De même la juxtaposition du nom propre "Socrate" et du verbe intransitif "parle" constitue une proposition. Pour isoler ce qu'il y a de commun aux prédicats "est impair" et à "chante", et ce qui les différencie des termes généraux "impair", "chanteur", "chantant", ...on pourrait dire que le prédicat contient, d'une manière explicite ou implicite, le verbe-copule "est". Le verbe "chante", selon cette vue, serait une forme condensée de "est chantant".

Ces considérations grammaticales ne nous retiendront pas plus longtemps. Tout ce que le logicien doit reconnaître se limite à ceci :

> les prédicats ne sont pas des noms ... Les prédicats et les termes singuliers sont ce que la prédication unit. (Quine 1970, 1975, p.48)

La différence entre sujets et prédicats ne réside donc pas uniquement dans une différence *sémantique*, comme c'est le cas pour la différence entre termes singuliers et termes généraux. Elle réside aussi dans une différence de *catégorie syntaxique*[1]. Dans la forme même du prédicat est inscrite *en creux* la possibilité de s'agréger à un nom pour former une proposition.

Un prédicat peut posséder plusieurs "creux", c'est-à-dire, plusieurs places. Par exemple le prédicat "aime" ou le prédicat d'appartenance "$\epsilon$" sont des prédicats à deux places : il faut, en effet, les combiner à deux termes singuliers ou à deux sujets[2] pour obtenir une proposition.

On pourrait penser que la notion même de prédicat est devenue superflue une fois que la notion de fonction propositionnelle a été introduite. Pourquoi ne dirait-on pas que la proposition :

(5.1) Pâris aime Hélène

[1] Deux expressions appartiennent à la même catégorie syntaxique si et seulement si elles sont intersubstituables *salva congruitate*, c'est-à-dire sans porter atteinte au caractère bien formé de l'assemblage. Le mot "table" figurant dans la phrase "Il table trop sur ses succès passés" appartient à la même catégorie que la locution "s'appuie" mais pas à la même catégorie que le mot "chaise".

[2] Par abus de langage, nous dirons que "Pâris" et "Hélène" sont *sujets* du prédicat à deux places "aime".

résulte de la substitution de termes singuliers "Pâris" et "Hélène" à $x$ et $y$ dans la fonction propositionnelle "$x$ aime $y$" ?

Cette identification des prédicats à n-place et de la fonction propositionnelle est commode et légitime, (cf. Hodges, 1977, p.154), mais nous lui préférerons cependant une autre analyse, celle qui consiste à traiter la fonction propositionnelle "$x$ aime $y$" comme une *expression composée*, que l'on obtient en unissant grâce à la *construction grammaticale* de *prédication* le prédicat "(−) aime (−)" aux termes singuliers $x$ et $y$ qui sont tous trois des expressions simples.

Nous suivons en cela Frege qui distingue la place vide "(-)" de la lettre "$x$" pour toutes les expressions fonctionnelles — qu'il s'agisse de fonctions propositionnelles comme "$x$ aime $y$" ou de fonctions de termes comme "$x+y$" — dans le passage suivant :

> Une expression fonctionnelle transporte toujours avec elle des places vides pour l'argument (au moins une place vide), que l'on indique habituellement — lors de l'analyse — par la lettre '$x$' qui remplit les places vides. Mais l'argument ne doit pas être considéré comme faisant partie de la fonction et la lettre '$x$' ne fait donc pas partie de l'expression fonctionnelle. [...] En conséquence, j'ai appelé non saturée la fonction [...] (Frege, cité par Sluga, 1980, p.141)

Ce n'est pas le lieu de se demander si à cette incomplétude *syntaxique* correspond une incomplétude dans le *réel* et dans la *pensée*[3].

Nous construirons la langue formelle du calcul des prédicats en définissant d'abord un *lexique*, c'est-à-dire un ensemble de *symboles simples* rangés en catégories syntaxiques et ensuite une *grammaire* dont les règles permettent de former des *symboles composés* (termes ou formules). Dans le chapitre 7, nous donnerons une sémantique à cette langue, c'est-à-dire que nous montrerons comment on peut rattacher les termes et formules à une réalité extérieure au langage, et traiter les symboles comme des *signifiants* renvoyant à des *signifiés*. Ceci nécessitera l'introduction des concepts de *structure*, de *domaine* et de *fonction d'interprétation*.

### 5.2.1. *Lexique*

Le lexique comprend :

1. un ensemble infini dénombrable de symboles de prédicat à 1-place $\{P_1(-), P_2(-), \ldots, P_k(-), \ldots\}$, à 2-places $\{P_1(-,-), P_2(-,-), \ldots, P_k(-,-),$

[3] Aristote croit à une correspondance de ce genre et aussi Spinoza qui distingue la *substance* qui est *en soi* et *conçue par soi* du *mode* qui est "dans la substance, par le moyen de laquelle il est aussi conçu". Le *mode* dont parle Spinoza, est ce qui correspond dans la réalité à un prédicat à une place. Le premier axiome de l'*Ethique* s'énonce "Tout ce qui est, est ou bien en soi, ou bien en autre chose". Les enjeux logiques et philosophiques de ce problème sont traités dans Cocchiarella, 1986.

$\ldots\}, \ldots,$ à $n$-places, ... Dans les exemples, nous nous autorisons l'usage de $F, G, H, K, R \ldots$ comme symboles de prédicats,

2. les variables propositionnelles $p, q, r, \ldots$ sont considérées comme des symboles de prédicats à 0 place,
3. un ensemble infini dénombrable de symboles de fonction ou foncteurs à 1-place $\{f_1(-), f_2(-), \ldots, f_k(-), \ldots\}$, à 2-places $\{f_1(-,-), f_2(-,-), \ldots, f_k(-,-), \ldots\}, \ldots,$ à $n$-places, .... Les foncteurs à 0-place sont des constantes d'individu $\{a_1, a_2, \ldots, a_n \ldots\}$. Par exemple "$(-)^2$" est un foncteur à une place, "$(-)\cdot(-)$" est un foncteur à deux places et "5!" est une constante d'individu,
4. un ensemble infini dénombrable de variables individuelles $\{u, v, w, x, y, z, u', v', w', x', y', z', u'', \ldots\}$,
5. les connecteurs $\{\neg, \wedge, \vee, \supset, \equiv\}$,
6. les quantificateurs $\{\forall u(\forall v \ldots), \exists u(\exists v \ldots)\}$.

*Remarques.*

1. Les symboles de prédicat (y compris de prédicat à 0-place tel que $p$) et les symboles de fonction sont des *symboles* dont l'interprétation *varie* avec la structure $\mathcal{S}$ (cf chapitre 7).
2. Les formules $p$ et $P(a_1, \ldots, a_n)$ sont des constantes prédicatives à 0-place, c'est-à-dire des propositions. La première est une proposition atomique *inanalysée*. La seconde est une proposition atomique *analysée* en sujet et prédicat. Cette différence existe dans la langue naturelle aussi ("Il pleut", "Albert est plus fort que Bernard").

### 5.2.2. *Grammaire*

Le calcul des prédicats avec symboles de fonctions comporte quatre constructions.

**RF1** La *formation des termes*[4] :

1. $x_1, \ldots, x_n \ldots$ et $f_1^0, \ldots, f_n^0 \ldots$ sont des termes.
2. Si $t_1, \ldots, t_n$ sont des termes, $f_i^n(t_1, \ldots, t_n)$ est un terme.

**RF2** La *prédication* qui consiste à insérer $n$ termes dans les $n$ places d'un prédicat à $n$ places :

Si $t_1, \ldots, t_n$ sont des termes, alors $p$ et $P_i^n(t_1, \ldots, t_n)$ sont des formules. En particulier, les symboles propositionnels $p$, $q$, $r$, ... sont

[4] Les exposants de $f$ ou de $P$ sont des marqueurs d'arité. $f^2$ indique que le symbole de fonction comporte deux places d'arguments.

des formules atomiques. On y adjoint parfois les deux constantes propositionnelles distinguées $\mathcal{T}$ et $\mathcal{F}$ qui expriment respectivement **V** et **F**.

**RF3** L'introduction de *connecteurs* :

Si $A$ est une formule, $\neg A$ l'est aussi[5].
Si $A$ et $B$ sont des formules, $(A \wedge B)$, $(A \vee B)$, $(A \supset B)$, $(A \equiv B)$ le sont aussi.

**RF4** La *quantification* qui consiste à préfixer une formule par un quantificateur :

Si $A$ est une formule, et si $x_i$ est une variable d'individu, $\forall x_i A$ et $\exists x_i A$ sont des formules.

### 5.2.3. *Portée et liaison d'un quantificateur*

La portée du quantificateur se limite à la *formule atomique* (et non au *terme*) qui la suit immédiatement, à moins qu'on n'étende sa portée par des parenthèses.

Dans

(5.2) $(\exists x)5x + 3 > x \supset x = x,$

les deux occurrences de $x$ indiquées par les flèches sont dans la *portée* du quantificateur et sont *liées* par ce dernier.

Dans

(5.3) $(\forall x)(5x + 3 > x \supset x = x),$

les quatre occurrences de $x$ indiquées par les flèches sont dans la *portée* du quantificateur et sont liées par lui.

Les quantificateurs en $x$ (respectivement en $y$, $z\ldots$) lient les variables $x$ (respectivement $y$, $z\ldots$) initialement *libres* qui tombent dans leur *portée*. Dans la formule 5.4, la *portée* du quantificateur $\forall x$ est l'espace entre parenthèses doubles, celle de $\exists y$ est l'espace marqué par les parenthèses simples, celle de $\exists x$ est uniquement $Kx$. Les *liaisons* sont indiquées par les flèches. Pour alléger l'écriture, nous omettrons les parenthèses entourant les arguments d'une fonction propositionnelle. Nous écrirons $Hxy$ au lieu de $H(x, y)$.

[5]Nous employons comme métavariables $A$, $B$ dans nos règles de formation au lieu de $X$ et $Y$ utilisés dans le calcul des propositions pour éviter qu'on ne croie que $X$, $Y$ sont en rapport avec les variables individuelles $x$, $y$.

(5.4) $\forall x((p \wedge Fx) \supset (Hyx \vee \exists x Kx)) \wedge \exists y(Gx \vee Fy)$

La formule ci-dessus est une fonction propositionnelle dans laquelle deux variables figurent à titre d'occurrence libre (l'$y$ de $Hyx$ et l'$x$ de $Gx$).

Une expression peut *tomber dans la portée* d'un quantificateur *sans être liée par* ce quantificateur. Par exemple, $p$ et $\exists x Kx$ de l'exemple 5.4 *tombent dans la portée* du quantificateur $\forall x$ sans contenir de variable libre que ce dernier puisse *lier*. La réciproque n'est pas vraie : une variable libre ne peut pas *être liée* par un quantificateur sans *tomber dans sa portée*.

Nous avons vu que les quantificateurs ont une action sur les expressions qui sont dans leur portée, même lorsque ces dernières ne sont pas liées par eux : la formule

(5.5) $\forall x$($x$ signe la pétition $\supset$ je serai surpris)

n'est *pas* logiquement équivalente à

(5.6) $\forall x$($x$ signe la pétition ) $\supset$ je serai surpris.

Le cas suivant, cependant, est déconcertant et pourrait passer pour un cas *dégénéré* de quantification :

(5.7) $\forall x(2 + 2 = 4 \supset 3 + 3 = 6)$.

On serait tenté de croire qu'une *proposition générale* de la forme $\forall x A$ ou $\exists x A$ *ne* peut être formée *qu'*en préfixant une formule contenant une variable libre au moyen d'un quantificateur ($\forall x$ ou $\exists x$). Ce n'est pas le cas. Si on rejetait comme mal formée une proposition de la forme

(5.8) $\forall x(p \supset q)$,

où le quantificateur fonctionne "à vide", on se priverait de certaines démonstrations indispensables (voir la remarque du paragraphe 5.3.3).

*Notion de terme libre pour une variable*

Soit $\Phi$ une formule du langage $\mathcal{L}$ décrit plus haut. Un terme $\tau$ est dit *libre pour* la variable $\nu$ dans $\Phi$ si $\nu$ ne figure pas libre en $\Phi$ dans le portée d'un quantificateur $\forall \mu$ où $\mu$ est une variable apparaissant dans $\tau$. En d'autres termes, $\tau$ peut être substitué à une occurrence libre de $\nu$ dans $\Phi$ à condition que cela ne provoque pas des interactions avec un quantificateur dans $\Phi$ (cf. Hamilton, 1978, p.55).

Par exemple, le terme $x + z$ est libre dans

(5.9) $Fx \supset \forall y Gxy$,

mais il ne l'est pas dans

(5.10) $\forall z \exists y Gxy \supset Fx$,

où $\Phi$ est $Gxy \supset Fx$,
$\nu$ est $x$
$\tau$ est $x + z$
parce que $\mu$, ici $z$, qui est une partie de $\tau$, est capturée par $\forall\mu$, c'est-à-dire par $\forall z$.

## 5.3. L'axiomatique du calcul des prédicats

Les lettres $\Phi, \Psi, \Xi$ sont des variables du métalangage. Elles prennent comme valeurs des *formules* du langage-objet. Les lettres $\mu, \nu$ sont des variables du métalangage qui prennent comme valeurs des *variables individuelles* du langage-objet telles que $x, y, z, u, v, \ldots$ La lettre $\tau$ est une variable du métalangage qui prend comme valeurs des *termes* du langage-objet.

On se donne les trois schémas d'axiome du calcul des propositions ci-dessous :

$A.1$ $\quad \Phi \supset (\Phi \supset \Psi)$

$A.2$ $\quad (\Phi \supset (\Psi \supset \Xi)) \supset ((\Phi \supset \Psi) \supset (\Phi \supset \Xi))$

$A.3$ $\quad (\neg\Psi \supset \neg\Phi) \supset ((\neg\Psi \supset \Phi) \supset \Psi)$

ainsi que les définitions de $\vee, \wedge, \equiv$ et la règle du *modus ponens*.

### 5.3.1. *Les nouveaux axiomes*

Aux trois schémas d'axiome du calcul des propositions, on ajoute les *deux schémas d'axiomes* suivants, propres au calcul des prédicats :

**A.4** $\forall\nu\ (\Psi \supset \Phi) \supset (\Psi \supset \forall\nu\Phi)$

où $\nu$ est une variable individuelle qui ne figure pas à titre d'occurrence libre dans $\Psi$. En revanche, $\Phi$ est n'importe quelle formule. Par exemple, les deux formules suivantes sont des instances de cet A.4 :

(5.11) $\qquad \forall x\ (p \supset Fx) \supset (p \supset \forall x Fx)$

(5.12) $\qquad \forall x\ (p \supset q) \supset (p \supset \forall x q)$.

Par contre, la formule 5.13 n'est pas une instance de A.4 :

(5.13) $\qquad \forall x\ (Fx \supset Gx) \supset (Fx \supset \forall x Gx)$

parce que $x$ est libre dans $Fx$. On trouve aisément une proposition fausse de langue naturelle qui a cette forme :

(5.14) $\qquad \forall x\ (x \text{ est pair} \supset x \text{ est pair}) \supset (x \text{ est pair} \supset \forall x\ x \text{ est pair})$.

**A.5** $\forall\nu\Phi \supset \Phi_{\tau/\nu}$

où $\nu$ est libre pour tout $\tau$ dans $\Phi$, c'est-à-dire où $\nu$ est libre et où son

substitut $\tau$ n'est pas capturé par un quantificateur de $\Phi$. Les deux formules suivantes sont des instances de A.5 :

(5.15) $\forall x Fx \supset Fy$

(5.16) $\forall x p \supset p$

En revanche, la formule suivante n'est pas une instance de A.5 car la variable $\tau$ dans $\Phi_{\tau/\nu}$ a été capturée par le quantificateur $\exists\tau$ présent dans $\Phi_{\tau/\nu}$ :

(5.17) $\forall x \underbrace{\exists y Rxy}_{\Phi} \supset \underbrace{\exists y Ryy}_{\Phi_{\tau/\nu}}.$

La phrase suivante de la langue naturelle qui a la forme de 5.17 est fausse :

(5.18) Si toute chose admet une chose plus grande qu'une autre, il y a une chose plus grande qu'elle même.

### 5.3.2. *La définition du quantificateur existentiel*

Le quantificateur universel est pris comme *terme primitif*. Le quantificateur existentiel, au contraire, est défini :

(5.19) $\exists x \Phi =_{def} \neg\forall x \neg\Phi.$

### 5.3.3. *Les règles d'inférence du calcul des prédicats*

**R.1** Le *modus ponens* est étendu aux formules ouvertes.

**R.2** La règle de *généralisation universelle* (G.U.) est introduite :

(5.20) $$\frac{\Phi}{\forall\nu\Phi}.$$

### 5.3.4. *La règle de généralisation universelle*

Une règle d'inférence est une transformation qui doit préserver la *valeur distinguée* des formules auxquelles on l'applique, en d'autres termes, c'est une opération qui transporte la vérité ou la validité des prémisses aux conclusions.

A strictement parler, seules les propositions peuvent être vraies ou fausses, seules les formes propositionnelles peuvent être valides ou falsifiables. La règle de généralisation universelle pose donc problème dans la mesure où sa prémisse — représentée par la marque place $\Phi$ — peut consister en une fonction propositionnelle. En effet les fonctions propositionnelles (2.1.4), contrairement aux propositions, n'ont pas de valeur de vérité déterminée.

Dans les mathématiques non formalisées cependant, on emploie souvent des fonctions propositionnelles comme des propositions dotées d'une valeur de vérité. En effet on sous-entend un quantificateur qui lie les variables libres. Par exemple, on traite l'*équation*

(5.21) $x + 3 = 5$

comme une proposition pourvue d'une valeur de vérité (le vrai). Le quantificateur sous-entendu est le quantificateur existentiel. De même, l'*identité*

(5.22) $x + y = y + x$

est traitée comme une proposition, mais cette fois, le quantificateur sous-entendu est le quantificateur universel.

Dans le contexte de la règle de généralisation universelle (G.U.)

(5.23) $\dfrac{\Phi}{\forall\nu\Phi}$ (G.U.)

dont

(5.24) $\dfrac{Px}{\forall x Px}$

est une instance, nous adopterons la deuxième interprétation. Prise dans cette interprétation la règle est manifestement correcte. Certes elle peut conduire à une conclusion fausse comme dans

(5.25) $\dfrac{Pair(x)}{\forall x\ Pair(x),}$

mais comme dans ce cas la prémisse est fausse (il n'est pas vrai que n'*importe quel* $x$ soit pair), la correction de la règle — à laquelle on demande seulement de *préserver* le vrai (ou le valide) — n'est pas mise en péril.

Dire que la règle est correcte, c'est dire qu'elle transporte le vrai (ou le valide) et qu'elle peut donc être appliquée à des prémisses — ici des fonctions propositionnelles — qui ont le statut de *thèses* (axiomes ou théorèmes de logique) ou de *prémisses* vraies extérieures à la logique. Peut-on l'appliquer aussi à des fonctions propositionnelles qui n'ont que le statut de *suppositions* ?

Avant de répondre à cette question, nous allons transposer l'instance particulière de G.U.

(5.26) $\dfrac{Px}{\forall x Px}$

en cet énoncé du métalangage dans lequel $Px$ reçoit explicitement le statut de supposition :

(5.27) De $Px$, on peut déduire $\forall x Px$

ce qui formellement s'écrit :

(5.28) $Px \vdash \forall x Px$.

Si maintenant nous appliquons le théorème de la déduction, nous obtenons :

(5.29) $\vdash Px \supset \forall x Px$.

En prenant comme précédemment la variable $x$ dans son acception de généralité, nous arrivons à

(5.30) $\vdash \forall x(Px \supset \forall x Px)$.

Or en vertu d'une loi de passage 5.52,

(5.31) $\forall x(Px \supset \forall x Px)$

est équivalent à

(5.32) $\exists x Px \supset \forall x Px$

qui est manifestement faux (si quelque $x$ est pair, tout $x$ est pair).

La règle de généralisation universelle est correcte et le théorème de la déduction du calcul des propositions est correcte. Comment la combinaison des deux règles isolément correctes peut-elle engendrer un raisonnement global incorrect ?

Stoll donne une explication claire de la genèse de l'erreur dans le passage suivant :

> [...] dans une démonstration qui contient $A(x)$ à titre de supposition (*assumption formula*), on n'a pas droit de généraliser $x$ (d'appliquer à $x$ la règle de généralisation universelle). En d'autres termes, $x$ est traité comme constante. Intuitivement, nous pouvons dire qu'une occurence libre d'une variable dans une supposition est employée pour dénoter un individu arbitraire sans doute, mais néanmoins fixé. Dans les mathématiques formalisées quand une variable $x$ est employée de cette manière on dit qu'elle reçoit l'*interprétation conditionnelle*. Au contraire, si $x$ figure libre dans une formule $A(x)$ qui est un *axiome* de la théorie considérée, alors $A(x)$ est censé signifier la même chose que $\forall x Ax$. Dans ce cas, on dit que $x$ reçoit l'*interprétation générique* ou "de généralité" (generality interpretation) (Stoll, 1961, p.392).

L'erreur dans la dérivation examinée plus haut tient en ceci : on est passé *subrepticement* de l'interprétation *générique* de la variable (en 5.26) à l'interprétation *conditionnelle* (en 5.28). Notons que certains auteurs réservent le mot "variable" pour l'interprétation générique et emploient le mot "paramètre" pour l'interprétation conditionnelle.

La règle G.U. est correcte pour autant que la variable généralisée $\nu$ s'il y en a une dans $\Phi$ soit prise dans l'interprétation générique (cf. chapitre 7, correction de l'axiomatique du calcul des prédicats).

La règle dérivée suivante — ou plus exactement le métathéorème — qui nous permettra d'utiliser la règle G.U. dans les dérivations à partir d'hypothèses sera également démontrée au chapitre 7.

(5.33) $$\frac{\Gamma \vdash \Phi}{\Gamma \vdash \forall \nu \Phi}$$

où $\nu$ n'apparait pas libre dans les formules de $\Gamma$.

Nous allons à présent formuler et démontrer la version du théorème de

la déduction qui peut être utilisée dans le calcul des prédicats et employée *conjointement* avec la règle G.U. et la règle 5.33. Avant celà, nous donnons quatre définitions.

### 5.3.5. *Définitions*

**Définition 5.1 : démonstration.** Une *démonstration* est une suite finie de formules $\Phi_1, \ldots, \Phi_n$ telle que chaque $\Phi_i$ est soit un axiome, soit dérivé à partir des membres antérieurs de la suite au moyen des règles M.P. et G.U.

**Définition 5.2 : théorème.** Si $\Phi$ est la dernière ligne d'une démonstration, $\Phi$ est un *théorème.*

**Définition 5.3 : dérivation à partir de prémisses membres de $\Gamma$.** Même définition que pour la démonstration, à cela près qu'on admet que des membres de $\Gamma$ figurent dans la suite.

**Définition 5.4 : $\Phi$ est dérivable de $\Gamma$ ($\Gamma \vdash \Phi$).** $\Phi$ est la dernière ligne d'une dérivation à partir des hypothèses figurant dans $\Gamma$.

### 5.3.6. *Théorème de la déduction et calcul des prédicats*

Le théorème de la déduction

$$(5.34) \quad \frac{\cdot\ \Gamma, X \vdash Y}{\Gamma \vdash X \supset Y}$$

peut néanmoins être étendu au calcul des prédicats, *à condition que l'on impose la restriction suivante* : la dérivation de $Y$ à partir de $\Gamma$ et de $X$ ne peut avoir fait appel à une application de la règle de généralisation universelle qui lie une variable qui était libre, soit dans la formule $X$, soit dans une formule appartenant à l'ensemble $\Gamma$.

Moyennant cette restriction, on peut étendre le théorème de la déduction au calcul des prédicats. Ce théorème se démontre par induction[6] sur la longueur de la démonstration. On suppose par hypothèse inductive que le théorème est vérifié pour toute $j$-ème formule antérieure à $i$, et on démontre qu'il est vérifié aussi pour la $i$-ème formule obtenue à partir d'une formule antérieure par application de la règle de généralisation.

Soit la ligne

$$\Gamma, X \vdash \forall x Y x,$$

[6] Il s'agit de l'*induction forte* qui s'écrit : $(\Phi 0 \wedge \forall k < m(\Phi k \supset \Phi m)) \supset \Phi m$ qu'il faut distinguer de l'*induction faible* qui s'écrit : $(\Phi 0 \wedge \forall m(\Phi m \supset \Phi S(m))) \supset \forall m \Phi m$ où $S$ désigne le successeur de $m$.

qui est la $i$-ème ligne et la ligne

$$\Gamma, X \vdash Y\boldsymbol{x},$$

qui est la ligne $j$ où $j < i$. Par hypothèse inductive, on peut appliquer le théorème à la $i$-ème ligne et passer de

$$\Gamma, X \vdash Y\boldsymbol{x}$$

à

$$\Gamma \vdash X \supset Y\boldsymbol{x}.$$

Nous allons montrer que l'on peut obtenir

$$\Gamma \vdash X \supset \forall \boldsymbol{x} Y\boldsymbol{x},$$

ce qui, sachant qu'on avait

$$\Gamma, X \vdash \forall \boldsymbol{x} Y\boldsymbol{x},$$

revient à étendre le théorème de la déduction aux formules obtenues grâce à la règle de généralisation universelle.

**Définition 5.5 : la dépendance.** Nous dirons qu'une formule $\Phi$ *dépend* d'une formule $\Psi$ ou d'un ensemble $\Gamma$ si $\Phi$ est identique à $\Psi$ ou à un membre de $\Gamma$, ou si $\Phi$ a été déduit par application des règles M.P. ou G.U. de quelques formules antérieures dont l'une au moins dépend de $\Psi$ ou de $\Gamma$, c'est-à-dire est identique à $\Psi$ ou à un membre de $\Gamma$, ou a été, à son tour, déduite par M.P. ou G.U., et *ainsi de suite*.

1^e^ **sous-cas :** $Y\boldsymbol{x}$ dépend de $X$ mais $\boldsymbol{x}$ n'est libre ni dans $X$, ni dans aucune formule qui en dépend. Comme $\boldsymbol{x}$ n'est pas libre dans $X$, on peut substituer $X$ à $\Psi$ et $Y\boldsymbol{x}$ à $\Phi$ dans l'axiome **A.4**. On obtient

$$\vdash \forall \boldsymbol{x}(X \supset Y\boldsymbol{x}) \supset (X \supset \forall \boldsymbol{x} Y\boldsymbol{x}),$$

et *a fortiori*

$$\Gamma \vdash \forall \boldsymbol{x}(X \supset Y\boldsymbol{x}) \supset (X \supset \forall \boldsymbol{x} Y\boldsymbol{x}).$$

Par hypothèse inductive, on a

$$\Gamma \vdash (X \supset Y\boldsymbol{x}),$$

et, par application de la généralisation universelle sur cette formule, on a

$$\Gamma \vdash \forall \boldsymbol{x}(X \supset Y\boldsymbol{x}).$$

On obtient donc, par *modus ponens* sur l'instance de l'axiome **A.4** et sur la formule précédente,

$$\Gamma \vdash (X \supset \forall \boldsymbol{x} Y\boldsymbol{x}).$$

$2^{e}$ **sous-cas :** $Yx$ ne dépend pas de $X$. Dans ce cas, il dépend de $\Gamma$ (ou ne dépend de rien, s'il est un axiome) et l'on peut écrire

$$\Gamma \vdash Yx.$$

On est donc ramené à un cas banal d'application de G.U. en dehors du contexte particulier du théorème de la déduction :

$$\Gamma \vdash \forall x Yx.$$

En substituant $\forall x Yx$ à $\Phi$ et $X$ à $\Psi$ dans l'axiome **A.3** du calcul des propositions, on obtient

$$\vdash \forall x Yx \supset (X \supset \forall x Yx),$$

et, par *modus ponens* :

$$\Gamma \vdash (X \supset \forall x Yx).$$

### 5.3.7. *Métathéorème de l'échange*

Si $C_A$ est une formule quelconque dont $A$ est un constituant (une sous-formule), si $C_B$ diffère de $C_A$ en cela que $A$ a été remplacé par $B$ quelquefois ou partout et si chaque variable libre de $A$ ou $B$ qui est aussi une variable libre de $C_A$ figure dans la liste $x_1, \ldots, x_n$, alors

(5.35) $\quad \vdash \forall x_1 \cdots \forall x_n \, (A \equiv B) \supset (C_A \equiv C_B).$

La règle d'échange pour le calcul des prédicats qui vient d'être énoncée est la deuxième partie de ce que Russell appelle le *principe d'extensionnalité*. La règle d'échange pour le calcul des propositions en est la première partie. Voici la formulation générale de ce principe couvrant les deux calculs :

Le principe d'extensionnalité se subdivise en deux parties :

1. "La valeur de vérité de toute fonction d'une proposition dépend uniquement de la valeur de vérité de l'argument, c'est-à-dire que si $p$ et $q$ sont toutes deux vraies ou toutes deux fausses, alors toute phrase contenant $p$ demeure vraie ou fausse, selon le cas, si l'on substitue $q$ à $p$." Par "fonction d'une proposition", Russell entend les propositions composées de propositions atomiques liées par des connecteurs vérifonctionnels $(\neg, \wedge, \ldots)$.
2. "La valeur de vérité de toute fonction d'une fonction propositionnelle ne dépend que de la seule extension de la fonction propositionnelle, c'est-à-dire que, si chaque fois que $\phi x$ est vrai, $\psi x$ est vrai, et réciproquement, alors tout énoncé portant sur la fonction $\phi$ demeure vrai ou faux, selon le cas, si l'on substitue $\psi$ à $\phi$" (Russell, 1940, 1969, p. 186).

*Démonstration de la règle d'échange.* La preuve se fait par induction sur le nombre $m$ de connecteurs et de quantificateurs de $C_A$.

$1^{\text{er}}$ **cas :** $A$ n'est remplacé nulle part par $B$.

On a alors

$$\vdash \forall x_1 \cdots \forall x_n (A \equiv B) \supset (C_A \equiv C_A),$$

qui est une instance de la tautologie suivante du calcul des propositions :

$$\vdash q \supset (p \supset p).$$

$2^{\text{ème}}$ **cas :** $C_A$ est identique à $A$.

Le théorème à démontrer est alors :

$$\vdash \forall x_1 \cdots \forall x_n (A \equiv B) \supset (A \equiv B).$$

On l'obtient par des applications répétées de l'axiome **A.4**.

$3^{\text{ème}}$ **cas :** $C_A$ est $\neg D$. Supposons que $C_B$ est de la forme $\neg D'$.

Par hypothèse d'induction, le théorème est supposé vérifié pour toute forme de complexité $k < m$; on a donc :

$$\vdash \forall x_1 \cdots \forall x_n (A \equiv B) \supset D \equiv D'.$$

Or, en vertu du calcul de propositions,

$$\vdash D \equiv D' \supset \neg D \equiv \neg D',$$

donc

$$\vdash \forall x_1 \cdots \forall x_n (A \equiv B) \supset \neg D \equiv \neg D'.$$

$4^{\text{ème}}$ **cas :** $C_A$ est $D \supset E$. Démonstration analogue.

$5^{\text{ème}}$ **cas :** $C_A$ est $\forall x D$. Supposons que $C_B$ est de la forme $\forall x D'$.

Par hypothèse d'induction,

$$\vdash \forall x_1 \cdots \forall x_n (A \equiv B) \supset D \equiv D'.$$

Comme on a imposé comme condition à l'application de la règle que les variables libres de $D$ et $D'$ *ne* sont *pas* libres dans $\forall x_1 \cdots \forall x_n (A \equiv B)$, on peut appliquer l'axiome **A.4**, ce qui donne

$$\vdash \forall x_1 \cdots \forall x_n (A \equiv B) \supset \forall x (D \equiv D'),$$

et, en vertu du théorème suivant du calcul des prédicats :

$$\vdash \forall x (D \equiv D') \supset (\forall x D \equiv \forall x D'),$$

on obtient

$$\vdash \forall x_1 \cdots \forall x_n (A \equiv B) \supset (\forall x D \equiv \forall x D').$$

### 5.3.8. *Règles dérivées*

**RD1** $\dfrac{\forall \nu \Phi}{\Phi}$ (Règle d'instantiation d'une universelle, I.U.)

*Démonstration.*

| | | | |
|---|---|---|---|
| (1) | $\forall x \Phi \supset \Phi_{\tau/\nu}$ | | Ax. 5 |
| (2) | $\forall x \Phi$ | | par hypothèse |
| (3) | | $\Phi_{\tau/\nu}$ | M.P.(1,2) |

**RD2** $\dfrac{\Phi}{\exists \nu \Phi}$ (Règle de généralisation existentielle, G.E.)

*Démonstration.*

| | | |
|---|---|---|
| (1) | $\forall \nu \neg \Phi \supset \neg \Phi_{\tau/\nu}$ | Ax. 5 |
| (2) | $\Phi_{\tau/\nu} \supset \neg \forall x \neg \Phi$ | (1) contraposition |
| (3) | $\Phi_{\tau/\nu}$ | par hypothèse |
| (4) | $\neg \forall \nu \neg \Phi$ | M.P. (2,3) |
| (5) | $\exists \nu \Phi$ | par la définition du quantificateur existentiel en termes du quantificateur universel |

Cette règle est facile à justifier *intuitivement* : si $\exists \nu \Phi$ est vrai, il y au moins un $\nu$ qui est $\Phi$, et que l'on peut désigner par $a$.

*Application*

| | | | |
|---|---|---|---|
| (5.36) | (1) | $\forall \nu(\Phi \supset \Psi) \supset (\Phi \supset \Psi)$ | Ax. 5 |
| | (2) | $\forall \nu(\Phi \supset \Psi)$ | Prémisse auxiliaire |
| | (3) | $\Phi \supset \Psi$ | M. P. 1, 2 |
| | (4) | $\forall \nu \Phi \supset \Phi$ | Ax. 5 |
| | (5) | $\forall \nu \Phi$ | Prémisse auxiliaire |
| | (6) | $\Phi$ | M. P. 4, 5 |
| | (7) | $\Psi$ | M. P. 3, 6 |
| | (8) | $\forall \nu \Psi$ | G.U. restrict. respectée |
| | (9) | $\forall \nu(\Phi \supset \Psi), \forall \nu \Phi \vdash \forall \nu \Psi$ | $(2,5 \vdash 8)$ |
| | (10) | $\forall \nu(\Phi \supset \Psi) \vdash \forall \nu \Phi \supset \forall \nu \Psi$ | Thm. de la déduction 9 |
| | (11) | $\vdash \forall \nu(\Phi \supset \Psi) \supset (\forall \nu \Phi \supset \forall \nu \Psi)$ | Thm. de la déduction 10 |

**RD3** $\dfrac{\exists \nu \Phi}{\Phi_{c/\nu}}$ (Règle d'instantiation d'une universelle, E.I.)

où $c$ est une constante nouvelle.

Dans les démonstrations informelles, après avoir établi une formule de la forme "Il existe un $\nu$ tel que $\Phi$", il est courant de poursuivre en disant : "appelons $c$ un tel $\nu$", où $c$ est une constante inédite appelée *paramètre*. On peut

introduire ce procédé dans un système axiomatique formalisé à titre de règle dérivée (dont on peut donc se passer), mais elle doit être accompagnée de restrictions : "aucune application de G.U. ne peut être faite en utilisant une variable qui est libre dans quelque $\exists\nu\Phi$ à laquelle la règle E.I. a été appliquée antérieurement" (Mendelson, 1965, p. 74). En outre, la dérivation doit pouvoir être réécrite de manière telle qu'elle ne contienne aucun des paramètres introduits par des applications de E.I. (Mendelson, *ibid.*, p. 73).

### 5.3.9. *Lois principales du calcul des prédicats*[7]

#### 5.3.9.1. *Lois de conversion des quantificateurs*

(5.37) $\exists\nu\Phi \equiv \neg\forall\nu\neg\Phi,$
(5.38) $\forall\nu\Phi \equiv \neg\exists\nu\neg\Phi,$
(5.39) $\neg\forall\nu\Phi \equiv \exists\nu\neg\Phi,$
(5.40) $\neg\exists\nu\Phi \equiv \forall\nu\neg\Phi.$

#### 5.3.9.2. *Lois de distribution des quantificateurs*

(5.41) $\forall\nu(\Phi \wedge \Psi) \equiv (\forall\nu\Phi \wedge \forall\nu\Psi)$
(5.42) $\exists\nu(\Phi \vee \Psi) \equiv (\exists\nu\Phi \vee \exists\nu\Psi)$

#### 5.3.9.3. *Lois de permutation des quantificateurs de même sorte*

(5.43) $\forall\nu\forall\mu\Phi \equiv \forall\mu\forall\nu\Phi,$
(5.44) $\exists\nu\exists\mu\Phi \equiv \exists\mu\exists\nu\Phi.$

#### 5.3.9.4. *Lois de réalphabétisation (renommage) des variables*

En vertu des théorèmes
(5.45) $\vdash \forall\nu\Phi \equiv \forall\mu\Phi,$

[7] La démonstration de ces lois se trouve notamment dans W. V. Quine, *Mathematical Logic*, 1940 Harvard U. P., 9 ème éd. revue 1981.

(5.46) $\vdash \exists\nu\Phi \equiv \exists\mu\Phi$,

et de la règle d'échange, on peut toujours renommer une variable liée et la variable du quantificateur au sein d'une formule. Par exemple, dans

(5.47) $\exists x(\forall x F x y \supset (G x \vee q))$,

on peut opérer deux renommages :

(5.48) $\exists u(\forall v F v y \supset (G u \vee q))$.

On peut renommer ("réalphabétiser") les variables liées, par exemple transformer $\exists x(Px \vee Qz)$ en $\exists y(Py \vee Qz)$, mais cette réalphabétisation ne peut entraîner la *capture* de variables libres : on ne peut pas remplacer $\exists y(Py \vee Qz)$ par $\exists z(Pz \vee Qz)$. On ne peut pas non plus réalphabétiser les variables libres; par exemple, on ne peut transformer $\exists x(Px \vee Qz)$ en $\exists x(Px \vee Qv)$. Le renommage est, dans certains cas, une transformation préparatoire à l'application des lois de passage, comme nous allons le voir.

### 5.3.9.5. *Lois de passage*

Si $\nu$ ne figure pas à titre d'occurrence libre dans $\Psi$, on a les lois suivantes dont une instance figure immédiatement en dessous de chaque formule :

(5.49) $\forall\nu(\Phi \wedge \Psi) \equiv (\forall\nu\Phi \wedge \Psi)$

$\forall x(Px \wedge q) \equiv (\forall x Px \wedge q)$,

(5.50) $\exists\nu(\Phi \wedge \Psi) \equiv (\exists\nu\Phi \wedge \Psi)$

$\exists x(Px \wedge q) \equiv (\exists x Px \wedge q)$,

(5.51) $\forall\nu(\Phi \vee \Psi) \equiv (\forall\nu\Phi \vee \Psi)$

$\forall x(Px \vee q) \equiv (\forall x Px \vee q)$,

(5.52) $\exists\nu(\Phi \vee \Psi) \equiv (\exists\nu\Phi \vee \Psi)$

$\exists x(Px \vee q) \equiv (\exists x Px \vee q)$,

(5.53) $\forall\nu(\Phi \supset \Psi) \equiv (\exists\nu\Phi \supset \Psi)$

$\forall x(Px \supset q) \equiv (\exists x Px \supset q)$,

(5.54) $\exists\nu(\Phi \supset \Psi) \equiv (\forall\nu\Phi \supset \Psi)$

$\exists x(Px \supset q) \equiv (\forall x Px \supset q)$,

(5.55) $\forall\nu(\Psi \supset \Phi) \equiv (\Psi \supset \forall\nu\Phi)$

$\forall x(q \supset Px) \equiv (q \supset \forall x Px)$,

(5.56) $\exists\nu(\Psi \supset \Phi) \equiv (\Psi \supset \exists\nu\Phi)$

(5.57) $\exists x(q \supset Px) \equiv (q \supset \exists x Px)$.

Il n'y a pas de loi de passage pour le biconditionnel. On peut cependant traiter $\forall\nu\Phi \equiv \Psi$ ou $\exists\nu(\Phi \equiv \Psi)$ en réduisant le biconditionnel grâce aux équivalences :

(5.58) $(p \equiv q) \equiv ((p \supset q) \wedge (q \supset p))$,

(5.59) $(p \equiv q) \equiv (pq \vee \overline{p}\ \overline{q})$.

Pour pouvoir appliquer une loi de passage, on doit parfois réalphabétiser. Par exemple, on ne peut appliquer la loi de passage 5.52 à la formule

(5.60) $\exists \boldsymbol{x} F\boldsymbol{x} \supset F\boldsymbol{x}$,

étant donné que $\nu$ figure à titre d'occurrence libre dans $\Psi$, c'est-à-dire ici, dans la deuxième occurrence de $F\boldsymbol{x}$. On peut remédier à cette situation en réalphabétisant $\exists \boldsymbol{x} F\boldsymbol{x}$ en $\exists y Fy$ :

(5.61) $\exists y Fy \supset F\boldsymbol{x}$.

On peut alors appliquer la loi 5.52, et on obtient :

(5.62) $\forall y(Fy \supset F\boldsymbol{x})$.

Dans certains cas, on a le choix, pour ce qui est de l'*ordre d'application* des lois de passage, entre plusieurs ordres. Si l'on part de la formule

(5.63) $\forall \boldsymbol{x} \exists y F\boldsymbol{x}y \supset \forall z \exists w Gzw$,

on peut appliquer les lois de passage aux quantificateurs en suivant l'ordre dans lequel ils apparaissent. On obtient alors successivement :

(5.64) $\exists \boldsymbol{x}(\exists y F\boldsymbol{x}y \supset \forall z \exists w Gzw)$,

(5.65) $\exists \boldsymbol{x} \forall y(F\boldsymbol{x}y \supset \forall z \exists w Gzw)$,

(5.66) $\exists \boldsymbol{x} \forall y \forall z(F\boldsymbol{x}y \supset \exists w Gzw)$,

et enfin

(5.67) $\exists \boldsymbol{x} \forall y \forall z \exists w(F\boldsymbol{x}y \supset Gzw)$.

On aurait pu appliquer les lois de passage d'abord au troisième quantificateur dans 5.63 ($\forall z$), puis au premier, puis au deuxième et enfin au quatrième de la manière suivante :

(5.68) $\forall z(\forall \boldsymbol{x} \exists y F\boldsymbol{x}y \supset \exists w Gzw)$,

(5.69) $\forall z \exists \boldsymbol{x}(\exists y F\boldsymbol{x}y \supset \exists w Gzw)$,

(5.70) $\forall z \exists \boldsymbol{x} \forall y(F\boldsymbol{x}y \supset \exists w Gzw)$,

(5.71) $\forall z \exists \boldsymbol{x} \forall y \exists w(F\boldsymbol{x}y \supset Gzw)$.

En revanche, *on n'aurait pas pu* commencer en appliquant une loi de passage au deuxième ou au quatrième quantificateur. En effet, l'application des lois de passage se fait en vertu du métathéorème de l'échange (section 5.3.7) qui stipule que *seuls* les *constituants* d'une formule peuvent être remplacés par des formules équivalentes. Il est incorrect de commencer en appliquant une loi de passage à $\exists y$ parce que cela exige de traiter $\exists y\, F\boldsymbol{x}y \supset \forall z \exists w\, Gzw$ comme un constituant de 5.63. Or il n'est pas un constituant de 5.63. Il n'apparaît nulle part dans l'arbre syntaxique de la formule en question (cf. figure 5.1).

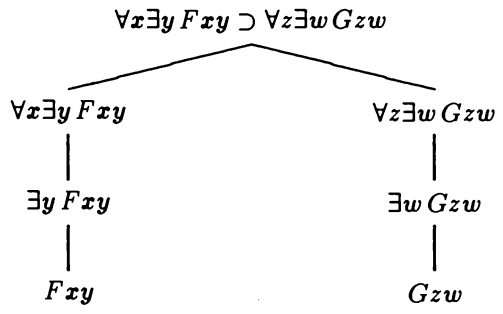



**Figure 5.1** *Arbre syntaxique de la formule 5.63*

### 5.3.10. *Applications*

*Remarque.* Les lettres $a, b, c$ sont des variables individuelles au même titre que $x, y, z$. Nos applications appartiennent toutes au calcul des prédicats (sans constantes individuelles ni foncteurs).

*Application 1*

(5.72) Si, de deux personnes, chacune est apparentée à une commune troisième, la première est apparentée à la deuxième. Chacun est apparenté à quelqu'un. Donc si John est apparenté à William, et William apparenté à Edith, alors John est apparenté à Edith (Kleene, 1967, 1972, p. 134).

*Représentation formelle*

| | |
|---|---|
| Hypothèse $H_1$ | $\forall x \forall y (\exists z (Rxz \wedge Ryz) \supset Rxy)$ |
| Hypothèse $H_2$ | $\forall x \exists y Rxy$ |
| Conclusion | $(Rab \wedge Rbc) \supset Rac$ |

Nous allons établir que la conclusion est dérivable de $H_1$ et $H_2$.

*Remarque.* Dans cet exercice réputé ardu (il est marqué d'un astérisque dans Kleene 1967), la difficulté réside dans le fait que les récurrences de termes singuliers dans la conclusion ne correspondent pas à celles de l'hypothèse $H_1$. Si l'on disposait d'une prémisse supplémentaire nous assurant que la relation $R$ est symétrique, il serait facile de faire en sorte que les récurrences de $a$ (respectivement $b$ et $c$) correspondent à celles des $x$ (respectivement $y$ et $z$), mais cette prémisse fait défaut. Il y a cependant moyen d'obtenir la permutation souhaitée ($Rbc$ au lieu de $Rcb$), grâce à des instantiations bien choisies de la ligne 8. *L'effort d'imagination requis par ce choix ne serait pas moindre en déduction naturelle.*

*Démonstration*[8].

| | | |
|---|---|---|
| (1) | $\forall x \forall y(\exists z(Rxz \wedge Ryz) \supset Rxy)$ | $H_1$ |
| (2) | $\forall x \exists y Rxy$ | $H_2$ |
| (3) | $\forall y(\exists z(Rcz \wedge Ryz) \supset Rcy)$ | I.U. sur $H_1$ |
| (4) | $\exists z(Rcz \wedge Rcz) \supset Rcc$ | I.U. sur 3 |
| (5) | $\exists z Rcz \supset Rcc$ | Echange et IDM sur 4 |
| (6) | $\exists z Rcz$ | I.U. et renom. sur $H_2$ |
| (7) | $Rcc$ | M.P. 5, 6 |
| (8) | $\forall x \forall y \forall z((Rxz \wedge Ryz) \supset Rxy)$ | Echange et $(\exists x \Phi \supset \Psi) \equiv (\forall x(\Phi \supset \Psi))$ sur $H_1$ |
| (9) | $(Rcc \wedge Rbc) \supset Rcb$ | I.U. sur 8 |
| (10) | $Rab \wedge Rbc$ | Prémisse auxiliaire |
| (11) | $(Rab \wedge Rbc) \supset Rbc$ | $\vdash (\Phi \wedge \Psi) \supset \Psi$ 10 |
| (12) | $Rbc$ | M.P. 11, 10 |
| (13) | $Rcc \wedge Rbc$ | $\vdash \Phi \supset (\Psi \supset (\Phi \wedge \Psi))$ 7, 12 M.P. |
| (14) | $Rcb$ | M.P. 9, 13 |
| (15) | $(Rab \wedge Rcb) \supset Rac$ | I.U. sur 8 |
| (16) | $(Rab \wedge Rbc) \supset Rab$ | $\vdash (\Phi \wedge \Psi) \supset \Phi$ |
| (17) | $Rab$ | M.P. 16, 10 |
| (18) | $Rab \wedge Rcb$ | $\vdash \Phi \supset (\Psi \supset (\Phi \wedge \Psi))$ 17, 14 M.P. |
| (19) | $Rac$ | M.P. 15, 18 |
| (20) | $Rab \wedge Rbc \vdash Rac$ | $H_1, H_2, 10 \vdash 19$ (voir app. p. 153) |
| (21) | $H_1, H_2 \vdash (Rab \wedge Rbc) \supset Rac$ | Théorème de la déduction |

[8] Voici une version élégante et sensiblement plus brève due à M A. Borgers :

| | | |
|---|---|---|
| (1) | $Rxz \wedge Ryz \supset Rxy$ | |
| (2) | $\exists z Rxz$ | |
| *Concl.* | $Rxz \wedge Ryz \supset Rxz$ | |
| (4) | $(Rxz \wedge Rxz) \supset Rxx$ | $1, x/y$ |
| (5) | $Rxz \supset Rxx$ | par $(p \wedge p) \equiv p$ |
| (6) | $\exists z Rxz \supset Rxx$ | par $\dfrac{Gz \supset P}{\exists x Gx \supset P}$ |
| (7) | $\exists z Rxz$ | (2) |
| (8) | $Rxx$ | M.P. 6,7 |
| (9) | $(Rxx \wedge Ryx) \supset Rxy$ | $1, x/z$ |
| (10) | $Ryx \supset Rxy$ | 8,9 et $\dfrac{(P \wedge Q) \supset R \quad P}{Q \supset R}$ |
| (11) | $Ryz \supset Rzy$ | 10 et $z/x$ |
| (12) | $Rxy \wedge Ryz \supset Rxy \wedge Rzy$ | 11 et $\dfrac{Q \supset S}{(P \wedge Q) \supset (P \wedge S)}$ |
| (13) | $Rxy \wedge Rzy \supset Rxz$ | 1, autres lettres |
| (14) | $Rxy \wedge Ryz \supset Rxz$ | 12,13 $\dfrac{P \supset Q \quad Q \supset R}{P \supset R}$ |

*Application 2*

(5.73) Si $x$ est père de $y$ et si $y$ est père de $z$, alors $x$ est grand-père de $z$. Tout le monde a un père. Donc tout le monde a un grand-père (Bergmann et Noll, 1977, p. 215).

*Représentation formelle*

$$\forall x \forall y \forall z((Pxy \wedge Pyz) \supset Gxz),$$
$$\forall x \exists y Pyx,$$
$$\text{donc } \forall x \exists y Gyx.$$

*Démonstration.* Appliquons deux fois successivement les règles I.U. et I.E. à la deuxième prémisse, nous obtenons :

(1) $\forall x \exists y\ Pyx, H_1$
(2) $\exists y\ Pya,$
(3) $Pba,$
(4) $\exists y\ Pyb,$
(5) $Pcb.$

Appliquons trois fois successivement la règle I.U. à la première prémisse, nous obtenons :

(6) $\forall x \forall y \forall z((Pxy \wedge Pyz) \supset Gxz), H_2$
(7) $\forall y \forall z((Pcy \wedge Pyz) \supset Gcz),$
(8) $\forall z((Pcb \wedge Pbz) \supset Gcz),$
(9) $(Pcb \wedge Pba) \supset Gca.$

En vertu de la loi suivante du calcul des propositions : $\Phi \supset (\Psi \supset (\Phi \wedge \Psi))$, on a

(10) $Pcb \supset (Pba \supset (Pcb \wedge Pba)),$

et par deux applications du *modus ponens* (10,5 et 10,3) on obtient

(11) $Pcb \wedge Pba.$

Puis, par *modus ponens* sur 9 et 11 :

(12) $Gca,$

par I.E. :

(13) $\exists y Gya,$

et par G.U. :

(14) $\forall x \exists y Gyx.$

---

*N.B.* On a appliqué la règle de substitution à des variables individuelles libres à plusieurs reprises, sachant qu'elles étaient liées par un quantificateur universel sous-entendu. On fait ainsi l'économie des règles U.I et U.G.

L'emploi de G.U. est licite étant donné que la variable $a$ n'est pas libre dans $\Sigma$, c'est-à-dire dans $\{H_1, H_2\}$.

*Application 3*

(5.74) $\vdash \forall x \forall y(Rxy \supset \neg Ryx) \supset \forall x \neg Rxx$.

*Démonstration.*

| | | |
|---|---|---|
| (1) | $\forall x \forall y(Rxy \supset \neg Ryx)$ | Prémisse auxiliaire |
| (2) | $\forall y(Rxy \supset \neg Ryx) \supset (Rxx \supset \neg Rxx)$ | dans A.5 |
| (3) | $(p \supset \neg p) \supset \neg p$ | théor. du calcul des prop. |
| (4) | $(Rxx \supset \neg Rxx) \supset \neg Rxx$ | 3, $Rxx/p$ |
| (5) | $((p \supset q) \wedge (q \supset r)) \supset (p \supset r)$ | syllog. du calcul des prop. |
| (6) | $((\forall y(Rxy \supset \neg Ryx) \supset (Rxx \supset \neg Rxx))$ $\wedge((Rxx \supset \neg Rxx) \supset \neg Rxx))$ $\supset (\forall y(Rxy \supset \neg Ryx) \supset \neg Rxx)$ | 5, $\forall y(Rxy \supset \neg Ryx)/p$, $Rxx \supset \neg Rxx/q$, $\neg Rxx/r$ |
| (7) | $\forall y(Rxy \supset \neg Ryx) \supset \neg Rxx$ | M.P. 6, conj.[9]de 2 et 4 |
| (8) | $\forall y(Rxy \supset \neg Ryx)$ | I.U. 1 |
| (9) | $\neg Rxx$ | M.P. 7,8 |
| (10) | $\forall x \neg Rxx$ | G.U. 11 |
| (11) | $\forall x \forall y(Rxy \supset \neg Ryx) \vdash \forall x \neg Rxx$ | $1 \vdash$ [10] |
| (12) | $\vdash \forall x \forall y(Rxy \supset \neg Ryx) \supset \forall x \neg Rxx$ | Th. de la déd. |

La difficulté réside dans le choix des axiomes et des substitutions de prédicats. *Ces difficultés disparaissent en déduction naturelle.*

### 5.3.11. *Raisonnement logique et démonstration géométrique*

Dans la *Critique de la Raison pure* (1781), Kant soutient que la démonstration de théorèmes de géométrie — par exemple l'affirmation que la somme des angles d'un triangle est égale à deux droits — requiert, en plus du *raisonnement logique*, le recours à l'intuition de l'espace :

> Que l'on donne à un philosophe le concept d'un triangle et qu'on le charge de trouver à sa manière quel peut être le rapport de ses angles avec l'angle droit. Tout ce qu'il a, c'est le concept d'une figure renfermée entre trois lignes droites et, dans cette figure, le concept d'un égal nombre d'angles. Il aura beau réfléchir, tant qu'il voudra, sur ce concept, il n'en fera sortir rien de nouveau. Il peut analyser et rendre clair le concept de la ligne droite, ou celui d'un angle, ou celui du nombre trois, mais non pas arriver à d'autres propriétés qui ne sont pas du tout contenues dans ces concepts. Mais que le géomètre s'empare de cette question. Il

[9] obtenue grâce à $\Phi \supset (\Psi \supset (\Phi \wedge \Psi))$ et à deux applications M.P. sur 2 et 4 respectivement.
[10] Introduction du *turnstile*, figure 3.14 (note 14, chap. 3).

commence aussitôt par construire un triangle. Sachant que deux angles droits pris ensemble valent autant que tous les angles contigus qui peuvent être tracés d'un point pris sur une ligne droite, il prolonge un côté de son triangle et obtient deux angles contigus dont la somme égale deux droits. Il partage ensuite l'angle interne en traçant une ligne parallèle au côté opposé du triangle et voit qu'il en résulte un angle contigu qui est égal à un angle externe, etc. De cette manière, il arrive, par une chaîne claire de raisonnements, toujours guidés par l'intuition, à la solution parfaitement claire et en même temps générale de la question (Kant, 1781, 1950, p. 495, cité par Beth, 1970, p. 47).

L'intuition de l'espace et le recours à un dessin peuvent faciliter la découverte des étapes du raisonnement, jouer un rôle heuristique ou didactique, mais jouent-ils un rôle strictement logique comme le croyait Kant ? On ne le croit plus aujourd'hui. Considérons le triangle suivant où les angles $\widehat{EAB}$ et $\widehat{ABC}$, que nous appelerons d et b, ainsi que $\widehat{FAC}$ et $\widehat{BCA}$, que nous appelerons $e$ et $c$, sont respectivement alternes internes.

Quand deux parallèles sont coupées par une sécante, les angles alternes-internes formés sont égaux.

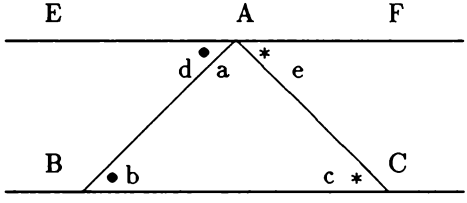



**Figure 5.2** *Somme des angles*

Représentons par les prédicats

$Txyz$, le fait que $x$, $y$, *et* $z$ *sont les angles d'un triangle,*

$Rxy$, le fait que $x$ *et* $y$ *sont dans un rapport d'alternes-internes,*

par le foncteur

$\circ(x,y,z)$, l'opération *d'additionner* $x$, $y$, *et* $z$,

et par la constante

$h$, la quantité *deux droits.*

A partir des axiomes de la géométrie notés $G$, on déduit par *la seule logique des prédicats*, et *sans* se préoccuper de l'interprétation des symboles $T(-,-,-)$, $R(-,-)$, $\circ(-,-,-)$ et $h$ :

(5.75) $G \vdash \forall x \forall y \forall z (Txyz \supset \exists u \exists v (Ryu \wedge Rzv \wedge \circ(x,u,v) = h))$,

(5.76) $G \vdash \forall x \forall y (Rxy \supset x = y)$.

En appliquant trois fois la règle d'instantiation d'universelle I.U. à 5.75, on obtient la formule suivante, où $a$, $b$, $c$ sont des *variables* d'individus et *non* des *constantes* désignant les angles particuliers du dessin :

(5.77) $G \vdash Tabc \supset \exists u \exists v(Rbu \wedge Rcv \wedge \circ(a,u,v) = h)$.

Par les lois de passage, on place les quantificateurs en tête :

(5.78) $G \vdash \exists u \exists v(Tabc \supset (Rbu \wedge Rcv \wedge \circ(a,u,v) = h))$.

Par la règle dérivée d'instantiation d'une existentielle I.E., on obtient la formule suivante, où $d$, $e$ sont des variables inédites :

(5.79) $G \vdash Tabc \supset (Rbd \wedge Rce \wedge \circ(a,d,e) = h)$.

On se donne à titre de prémisse $Tabc$ et on obtient, par *modus ponens*, sous réserve de la prémisse

(5.80) $G, Tabc \vdash Rbd \wedge Rce \wedge \circ(a,d,e) = h$,

et par application d'un théorème trivial du calcul des propositions ($\vdash (p \wedge q) \supset p$), on a que $p \wedge q \vdash p$ :

(5.81) $G, Tabc \vdash Rbd, Rce, \circ(a,d,e) = h$.

En appliquant deux fois la règle I.U. à 5.76, on obtient :

(5.82) $G \vdash Rbd \supset b = d$

(5.83) $G \vdash Rce \supset c = e$,

et, par deux applications du *modus ponens*, dans 5.82, 5.83 et 5.81

(5.84) $G, Tabc \vdash b = d, c = e$.

En appliquant la logique de l'identité (cf. 1.5.2), qui permet de substituer l'un à l'autre les membres d'une identité, nous remplaçons $d$ par $b$ et $e$ par $c$ dans

(5.85) $G, Tabc \vdash \circ(a,d,e) = h$,

pour obtenir

(5.86) $G, Tabc \vdash \circ(a,b,c) = h$.

Nous appliquons deux fois le théorème de la déduction. *C'est permis ici puisque l'on n'a pas fait appel à la règle G.U. jusqu'à présent.* On obtient :

(5.87) $\vdash G \supset (Tabc \supset \circ(a,b,c) = h)$.

A présent, nous pouvons appliquer G.U. à cette formule *qui est un théorème de logique*, et nous obtenons par *modus ponens* sur 5.87 et $G$ le *théorème de géométrie* :

(5.88) $\forall x \forall y \forall z[Txyz \supset \circ(x,y,z) = h]$.

Pour Kant, en géométrie, on doit recourir à l'intuition :

> [...], [la géométrie] ne peut rien faire avec le simple concept, mais se hâte de recourir à l'intuition où elle considère le concept *in concreto*, non pas cependant empiriquement, mais simplement dans une intuition qu'elle a

> représentée *a priori* c'est-à-dire qu'elle a construite et dans laquelle ce qui résulte des conditions générales de la construction doit s'appliquer aussi d'une manière générale à l'objet (*Objecte*) du concept construit (Kant, *Ibid.*, p. 495).

En d'autres termes, le géomètre doit contempler un triangle concret, puis faire abstraction de ses particularités.

La logique des prédicats avec identité permet de faire l'économie de ce détour par l'intuition de l'espace. Il est vrai que le logicien a dû *particulariser* ou *instancier* — grâce à la règle I.U. — puis *généraliser* — grâce à la règle G.U. — *mais nulle part il n'a besoin de l'intuition de l'espace.* La représentation spatiale, le dessin, a une valeur *heuristique* : il aide à *trouver* la démonstration logique, mais celle-ci une fois construite devient totalement indépendante du dessin. L'évidence de la démonstration formelle ne dépend pas de ce dernier.

## 5.4. La méthode de déduction naturelle

### 5.4.1. *Introduction*

Pour obtenir un système de déduction naturelle pour le calcul des prédicats, il suffit d'ajouter aux règles d'introduction et d'élimination des connecteurs les règles d'introduction et d'élimination des quantificateurs. Ces quatre règles ont déjà été invoquées dans l'exposé de la méthode axiomatique ($\forall$**I** appelée ici G.U., $\exists$**I** appelée G.E., $\forall$**E** appelée ici I.U. et $\exists$**E** appelée I.E.). Nous adopterons ici la version de ces règles présentée par Quine comme l'a fait Massey.

### 5.4.2. *Introduction et élimination de quantificateurs*

Il y a deux quantificateurs ($\forall x$ et $\exists x$) que l'on peut soit éliminer, soit introduire. Il y aura donc *quatre* règles. Dans chaque inférence autorisée par ces règles, l'une des deux lignes de l'inférence contient un quantificateur de plus que l'autre. On l'appelle *ligne générique.* L'autre ligne — celle qui contient un quantificateur de moins, voire pas de quantificateur du tout, — on l'appelle *ligne instanciée.* En cas *d'élimination* de quantificateur, la ligne instanciée *suit* la ligne générique. En cas d'*introduction* de quantificateur au contraire, elle *précède* celle-ci. On appelle *variable d'instantiation* la variable libre de la ligne instanciée, qui est liée dans la ligne générique.

Deux règles sur quatre sont faciles à justifier. Elles correspondent, en effet, à des conditionnels valides. Ce sont respectivement la règle d'*instantiation d'une universelle* (I.U.) qui permet *d'éliminer* le quantificateur *universel* :

- ligne générique : $\forall x \Phi x$
- ligne instanciée : donc $\Phi y$

- variable d'instantiation : $y$,

et la règle de *généralisation existentielle* (G.E.) qui permet *d'introduire* le quantificateur *existentiel* :

- ligne instanciée : $\Phi y$
- variable d'instantiation : $y$
- ligne générique : donc $\exists x \Phi x$.

*Démonstration de la correction de I.U.*. Pour montrer que la conclusion $\Phi y$ est impliquée par la prémisse, $\forall x \Phi x$, il suffit de prouver que le conditionnel $\forall x \Phi x \supset \Phi y$ est valide. Ce conditionnel est un schéma ouvert. On a donc : $\models \forall x \Phi x \supset \Phi y$ si et seulement si $\models \forall y(\forall x \Phi x \supset \Phi y)$ (cf. 5.3.5). Or on sait que le dernier schéma cité est équivalent à $\forall x \Phi x \supset \forall y \Phi y$ par les lois de passage et que $\forall x \Phi x \supset \forall y \Phi y$ est équivalent à $\forall x \Phi x \supset \forall x \Phi x$ par renommage des variables. Comme ce dernier est valide — c'est une instance de la tautologie $p \supset p$ — on peut conclure que $\forall x \Phi x \supset \Phi y$ est valide.

La démonstration de G.E. est analogue à celle de I.U.

Contrairement aux règles I.U. et G.E., les règles I.E. et G.U. ne correspondent pas à des théorèmes. Montrons-le pour I.E. en procédant par l'absurde :

(5.89) $\exists x \Phi x \supset \Phi y \equiv \forall y(\exists x \Phi x \supset \Phi y)$,

(5.90) $\forall y(\exists x \Phi x \supset \Phi y) \equiv (\exists x \Phi x \supset \forall y \Phi y)$,

(5.91) $\exists x \Phi x \supset \forall y \Phi y \equiv (\exists x \Phi x \supset \forall x \Phi x)$.

Or il est clair que

(5.92) $\not\vdash \exists x \Phi x \supset \forall x \Phi x$.

On ne peut pas inférer de "quelqu'un parle" l'énoncé "tous parlent".

Dans la présentation axiomatique du calcul des prédicats, on exigeait que G.U. ait comme prémisse un axiome ou un théorème, ce qui n'est plus le cas en déduction naturelle. La justification de G.U. prouvant que celle-ci est correcte (*sound*) n'est donc plus valable dans ce nouveau contexte. Quant à la règle dérivée I.E. de la méthode axiomatique, la variable d'instantiation qu'elle introduit n'est pas vraiment une variable, mais plutôt un nom propre créé pour la circonstance que nous appelons "paramètre". ("Supposons que nous ayons démontré une formule de la forme $\exists x \Phi x$. Il y a au moins un objet qui est $\Phi$, appelons-le $b$"). Dans le contexte de la déduction naturelle, au contraire, on a une variable $y$ dans la conclusion. *La justification de I.E. qui valait pour la méthode axiomatique ne vaut donc plus ici.* Une autre justification sera nécessaire.

### 5.4.3. *La généralisation universelle dans le raisonnement informel*

Fine observe que l'application de la règle G.U. à une prémisse qui n'est pas un axiome ou un théorème de logique, est pratique courante chez le

mathématicien :

> Supposons que je désire établir que tous les nombres naturels admettent une décomposition [unique] en facteurs premiers. Je prends un nombre naturel arbitraire $x$, j'établis qu'il a une décomposition unique en facteurs premiers, et je conclus de là que tous les nombres naturels ont une décomposition en facteurs premiers... (Fine, 1985, p. 137)

C'est exactement la démarche par laquelle Feferman démontre le théorème fondamental de l'arithmétique dû à Gauss (1802).

**Théorème 5.1.** Factorisation des entiers en nombres premiers.

*Démonstration.*

> [ ...] Considérons n'importe quel $a$ [entier plus grand que 1]. Si $a$ est premier, nous avons fini. S'il ne l'est pas, $a = a_1 \cdot a_2$ pour quelque $a_1, a_2$ tel que $1 < a_1 < a$ et $1 < a_2 < a$. Si nous supposons que le résultat vaut pour tout $b < a$, nous pouvons conclure qu'il vaut pour $a$. (Feferman 1964, p. 139)

*Esquisse de formalisation.*

Convenons d'utiliser $P$ pour "premier", $F$ pour "factorisable" et $I$ (l'initiale de "*integer*") pour "entier naturel différent de 0 et de 1".

| | | | | |
|---|---|---|---|---|
| (1) | $Ia$ | $\wedge$ | $a > 1$ | Prémisse auxiliaire ($P$) |
| (2) | $Pa$ | $\vee$ | $\neg Pa$ | théorème $\vdash \Phi \vee \neg\Phi$ |
| | $\vdots$ | | $\vdots$ | |
| $(n-1)$ | $Pa \supset Fa$ | | $\neg Pa \supset Fa$ | |
| $(n)$ | $Fa$ | $\vee$ | $Fa$ | Dilemme constructif 2, $n-1$, $n$ |
| $(n+1)$ | $Fa$ | | | Idemp. $n+1$ |
| $(n+2)$ | $(Ia \wedge a > 1) \supset Fa$ | | | Décharge de $P$ |
| $(n+3)$ | $\forall x((Ix \wedge x > 1) \supset Fx)$ | | | G.U. $n+2$ |

Notre présentation n'est qu'une reconstruction partielle de la démonstration informelle de Feferman. Une présentation formelle complète exigerait un recours à l'induction mathématique dans le langage-objet[11], ce qui sort du cadre de ce chapitre (cf. chapitre 9). Toute schématique qu'elle est, cette présentation suffit à illustrer le recours implicite à la généralisation universelle dans le raisonnement informel.

Dans la démonstration informelle de Feferman, on ne trouve que les étapes (1) à $(n+2)$, d'où découle le théorème d'arithmétique de la ligne $(n+3)$.

[11] Jusqu'ici nous n'avons utilisé l'induction mathématique que dans le métalangage.

Mais entre ce théorème qui concerne un entier quelconque $a$ et le théorème fondamental de l'arithmétique formulé en $(n+3)$ qui parle de *tous* les entiers, il y a un hiatus que vient combler un usage implicite de la règle G.U.

On voit assez facilement les exigences que G.U. doit satisfaire pour être une règle correcte dans un raisonnement informel : l'objet représenté par la variable $x$ doit vraiment être un objet *quelconque* et la prémisse de la règle ne doit pas être une *pure supposition*. Ces conditions sont satisfaites ici. On notera que la variable $a$ en (1) est prise dans l'interprétation conditionnelle — c'est un paramètre —, mais que en $(n+2)$ elle est prise dans l'interprétation générique. Les difficultés surgissent quand on tente d'exprimer ces exigences dans un système formel. Une manière d'esquiver ces difficultés est de montrer que *sous certaines conditions* une dérivation $\delta$ contenant des applications de G.U. peut être réécrite en une dérivation $\delta'$ qui n'en contient plus, quoiqu'elle ait la même conclusion que $\delta$.

### 5.4.4. *L'usage de règles non valides*

Il y a des cas où la chaîne d'inférences

(5.93) Quelque chose est $F$,
donc $x$ est $F$,
$x$ est $F$,
donc tout est $F$

conduit du vrai au vrai. Ce sera le cas, par exemple, si l'on travaille avec un modèle dont le domaine est un ensemble à un élément.

Une telle limitation est intolérable, mais on peut songer à des restrictions moins drastiques dont l'effet serait de conférer aux applications des règles G.U. et I.E. le statut d'*artifices* qui se révèlent inoffensifs pour autant que la déduction soit conduite à son terme, ce qui suppose le définition du concept de *dérivation achevée* (cf. 5.4.6). On n'exigera plus que la dérivation soit correcte (c'est-à-dire bonne conductrice du vrai) ligne par ligne, mais seulement qu'elle soit correcte prise comme un tout.

Selon la conception traditionnelle, la déduction est comparable à une chaine. Or la solidité d'une chaîne se mesure à la solidité du plus fragile de ses maillons. Dans la version de déduction naturelle exposée ici — à savoir celle de Quine —, une dérivation peut être correcte une fois *achevée*, tout en contenant des étapes qui sont incorrectes *quand on les considère isolément* — en d'autres termes, elle peut être correcte prise comme un *tout* et incorrecte dans ses *parties*. La métaphore de la chaîne ne rend plus justice à cette conception de la déduction. Il faut la remplacer par la métaphore de la voûte dont les constituants se soutiennent mutuellement. Nous avons choisi cette version de la déduction naturelle parce que la preuve de la correction de ses règles est particulièrement aisée. Nous devons cependant reconnaître que l'acceptation

de contraintes *globales* plutôt que locales est un défaut. Elle fait perdre aux preuves la propriété de *modularité* (cf. 5.4.7).

### 5.4.5. *Les restrictions imposées aux règles G.U. et I.E.*

Quine a formulé les deux restrictions suivantes :

> Les variables d'instantiation de I.E. et de G.U. doivent être différentes lors de chaque application, et la variable d'instantiation lors de chaque application doit être alphabétiquement postérieure à toutes les variables libres de la ligne générique de l'application en question (Quine, 1950, 1972, p. 218).

Pour faciliter le contrôle du respect de ces restrictions, Quine introduit un procédé : la signalisation. *Signaliser*[12] une application de I.E. ou de G.U., c'est mentionner la variable d'instantiation à droite de la ligne que produit cette application : dans le cas d'une application de G.U., la variable sera donc à droite de la ligne *quantifiée* alors que, dans le cas de I.E., elle sera à droite de la ligne *instanciée*, mais dans les deux cas elle sera à droite de la ligne *déduite*.

On peut reformuler ainsi les restrictions portant sur I.E. et G.U. :

> Aucune variable ne peut être deux fois l'objet d'une signalisation et la variable signalisée doit être alphabétiquement postérieure à toutes les variables libres de la ligne générique (Quine, *Ibid.*, p. 219).

#### 5.4.5.1. *Exemple de déduction incorrecte violant la première exigence*

(5.94)
(1) $\exists x F x$
(2) $F y \quad y$
(3) $\forall z F z \quad y$

Le passage de (1) à (2) est une application de la règle I.E., la variable d'instantiation est $y$, la ligne générique est (1). On place $y$ à droite de la ligne (2) car la ligne (2) est la ligne déduite.

Le passage de (2) à (3) est une application de la règle G.U., la variable d'instantiation est à nouveau $y$, car dans le cas des règles d'introduction de quantificateur, la ligne *instanciée précède* la ligne générique. Néanmoins, la ligne *déduite suit* — (3) est déduite de (2) — et dès lors, la variable d'instantiation doit être signalisée à droite de (3).

Les variables d'instantiation de I.E. et G.U. ne diffèrent pas. La première contrainte n'a donc pas été respectée. Le non-respect de cette contrainte conduit à un raisonnement manifestement non valide, à savoir "Quelque chose est F donc tout est F".

[12] Ce néologisme a été introduit pour traduire le mot anglais "flagging".

5.4.5.2. *Exemple de déduction incorrecte violant la deuxième exigence*

(5.95)

| | | | |
|---|---|---|---|
| (1) | $\forall x \exists w$ | $Fxw$ | |
| (2) | $\exists w$ | $Fvw$ | |
| (3) | | $Fvy$ | $y$ |
| (4) | $\forall x$ | $Fxy$ | $v$ |
| (5) | $\exists w \forall x$ | $Fxw$ | |

Le passage de (2) à (3) dans 5.95 est une application de la règle I.E. La variable d'instantiation $y$ écrite à droite de la ligne déduite (3) est alphabétiquement postérieure à la variable libre de la ligne générique, c'est-à-dire à la variable $v$ qui est libre dans la ligne générique (2).

Le passage de (3) à (4) est une application de la règle G.U. La variable d'instantiation est $v$ puisque la ligne (3) est la ligne instanciée et (4) la ligne générique. Comme (4) est déduite de (3), c'est à droite de la ligne (4) qu'on écrit la variable d'instantiation.

La variable signalisée est donc différente lors des applications successives de I.E. et de G.U. En revanche, la deuxième contrainte n'est pas respectée. En effet, la variable d'instantiation de G.U. qui assure le passage de (3) à (4) est $v$. Or dans la ligne générique de cette G.U., à savoir (4), on trouve la variable libre $y$. La variable d'instantiation $v$ de G.U. n'est donc pas alphabétiquement postérieure à la variable libre de la ligne générique dont elle est issue.

On pourrait penser qu'il suffit, pour corriger cette erreur, de remplacer la variable d'instantiation $v$ par $z$ lors de l'application de I.U. qui permet le passage de (1) à (2). Ce remplacement est permis puisque la règle I.U. n'est soumise à aucune contrainte, si ce n'est la contrainte gouvernant la substitution en général[13]. On obtient alors la dérivation 5.96 :

(5.96)

| | | | |
|---|---|---|---|
| (1) | $\forall x \exists w$ | $Fxw$ | |
| (2) | $\exists w$ | $Fzw$ | |
| (3) | | $Fzy$ | $y$ |
| (4) | $\forall x$ | $Fxy$ | $z$ |
| (5) | $\exists w \forall x$ | $Fxw$ | |

Cette dérivation est tout aussi fautive que la première, la faute n'a été que déplacée. Elle se situe à présent lors de l'application de I.E. En effet, la variable d'instantiation $y$ n'est pas, comme elle aurait dû l'être, alphabétiquement postérieure à la variable libre $z$ de la ligne générique (2).

Il n'est donc pas possible de contourner la deuxième exigence et de justifier le raisonnement non valide "tous les nombres naturels ont un successeur immédiat donc il y a un successeur immédiat (le même) à tous les nombres naturels".

[13] Soit $\nu_i$ une variable d'instantiation. Soient $\Phi$ et $\Phi_i$ deux formules — $\Phi_i$ étant une instance de $\Phi$ — la variable $\nu_i$ doit être libre dans $\Phi_i$ à tous les endroits où $\nu$ est libre dans $\Phi$.

Aux règles d'introduction et d'élimination des quantificateurs empruntées à Quine, nous ajouterons :

1. les règles d'introduction et d'élimination des connecteurs (2.6.3),
2. les règles accessoires du calcul des propositions (2.6.7),
3. les lois de De Morgan pour les quantificateurs,
4. les lois de distribution des quantificateurs,
5. les lois de permutation des quantificateurs de même quantité,
6. les lois de renommage,
7. les lois de passage,
8. la règle d'échange, implicitement mentionnée dans 2, qui permet d'utiliser pleinement les lois 3 à 7.

### 5.4.6. *Justification de I.E. et de G.U.*

Les règles I.U. et G.E. ne présentent pas de difficulté puisqu'elles correspondent à des schémas quantificationnels valides. Pour justifier les règles I.E. et G.U., on montre que dans tous les cas où les contraintes auxquelles elles sont soumises sont respectées, on peut reconstruire une dérivation dans laquelle les deux règles suspectes sont remplacées par un banal *modus ponens* combiné avec l'introduction d'une *prémisse auxiliaire* que l'on pourra réduire à une *tautologie*. Les deux règles peuvent donc être comparées à un échafaudage que l'on défait après avoir construit un édifice avec son aide. La réciproque est vraie : quand les contraintes ne sont pas respectées, on ne parvient pas à se débarrasser de ces prémisses auxiliaires et dès lors on est condamné soit à conserver dans la démonstration des applications de I.E. et G.U., ce qui *disqualifie* la démonstration, soit à conserver les prémisses auxiliaires, ce qui revient à remplacer une démonstration incorrecte d'une proposition par une démonstration correcte, mais qui démontre une *autre proposition* que la proposition que l'on cherchait à démontrer.

Pour démontrer de manière tout à fait générale que les règles I.E. et G.U. ne sont finalement que des artifices pour abréger les démonstrations, on a recours à une induction mathématique et celle-ci fait intervenir le concept de "dérivation achevée" qu'il faut commencer par définir. Une dérivation est *achevée* ou *finie* "si aucune variable faisant l'objet d'une signalisation n'est libre dans sa dernière ligne ou dans une prémisse quelconque dont dépend cette dernière ligne" (Quine). Ce n'est que pour les dérivations achevées que l'on affirme que les prémisses de la dérivation impliquent la conclusion, c'est-à-dire que l'on affirme que la conclusion est vraie si les prémisses le sont. Une dérivation achevée peut, dans certains cas, être continuée. Ce sera le cas si elle contient des prémisses non résorbées. Si on élimine toutes les prémisses en appliquant la règle d'introduction du conditionnel, c'est-à-dire si l'on traite toutes

les prémisses comme des prémisses auxiliaires, on transforme la dérivation en *démonstration* et la conclusion en *théorème*. Un théorème est, en effet, défini comme une conclusion découlant de la classe nulle des prémisses ou comme une prémisse à "zéro conclusion".

On a donc les trois possibilités présentées en italique aux extrémités de l'arbre renversé à la figure 5.3.

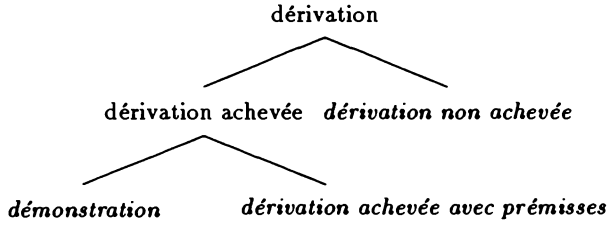



**Figure 5.3** *Trois sortes de dérivation*

Pour justifier I.E. et G.U., on démontre l'induction mathématique suivante :

**Base :** Une dérivation achevée avec 0 variable signalisée est correcte. En effet, si elle a 0 variable signalisée, elle n'utilise jamais I.E. et G.U.

**Induction :** Si les dérivations achevées avec $n$ variables signalisées sont correctes, alors les dérivations achevées avec $n + 1$ variables signalisées le sont aussi.

Considérons la dérivation suivante qui est achevée et qui contient $n+1$ variables signalisées (en fait elle en contient 2). Supposons par hypothèse inductive que toutes les dérivations achevées qui contiennent $n$ variables signalisées sont correctes. Essayons de réduire la dérivation 5.97 à une dérivation à $n$ variables signalisées. Nous allons éliminer la variable signalisée la plus proche du début de l'alphabet (dans le cas présent, $v$).

(5.97)

| | | | |
|---|---|---|---|
| (1) | $\exists x \forall y A x y$ | | |
| | | ↓ I.E. | |
| (2) | $\forall y A z y$ | | $z$ |
| | | ↓ I.U. | |
| (3) | $A z v$ | | |
| | | ↓ G.E. | |
| (4) | $\exists x A x v$ | | |
| | | ↓ G.U. | |
| (5) | $\forall y \exists x A x y$ | | $v$ |

Pour y parvenir, nous ajoutons la prémisse auxiliaire $\exists x A x v \supset \forall y \exists x A x y$. Cette prémisse auxiliaire n'est autre que le conditionnel formé en prenant pour antécédent la ligne à laquelle on applique la règle G.U. et pour conséquent la ligne qui résulte de l'application de la dite règle. Appelons (0) cette prémisse auxiliaire. Grâce à elle, on peut tirer (5) de (0) et (4) à l'aide d'un banal

*modus ponens*. Dès lors on se passe de G.U. et on est libéré de l'obligation de signaliser la variable $v$. On a donc réduit la dérivation à une dérivation à $n$ variables signalisées.

Sachant que la variable signalisée $v$ était par hypothèse la variable la plus proche du début de l'alphabet parmi toutes les variables signalisées de la dérivation en question, et sachant que $v$ était alphabétiquement postérieure à toutes les variables libres de la ligne générique (5) en vertu du fait que la dérivation respecte les contraintes, on peut conclure que (0) ne contient aucune variable libre signalisée. L'adjonction de la ligne (0) prise comme prémisse auxiliaire n'a donc pas fait perdre à la dérivation son statut de dérivation achevée comme cela aurait été le cas si la nouvelle prémisse avait contenu une variable signalisée libre. En revanche, elle a rendu superflue la $n+1^{\text{ème}}$ variable signalisée. La dérivation allongée grâce à (0) est une dérivation achevée ayant $n$ variables signalisées. En vertu de l'hypothèse inductive, cette dérivation allongée d'une ligne est correcte. En d'autres termes, ses prémisses (0) et (1) impliquent la conclusion (5).

Cela revient à dire que le conditionnel (6) est valide :

$$(6) \qquad [\underbrace{(\exists x A x v \supset \forall y \exists x A x y)}_{(0)} \wedge (\exists x \forall y A x y)] \supset \forall y \exists x A x y.$$

Ce conditionnel (6) est une formule ouverte. Or toute formule ouverte valide est équivalent à sa clôture par un quantificateur universel :

(5.98) $\models \Phi x$ est équivalent à $\models \forall x \Phi x$.

On obtient donc :

$$(7) \qquad \forall v(((\exists x A x v \supset \forall y \exists x A x y) \wedge (\exists x \forall y A x y)) \supset \forall y \exists x A x y).$$

Comme $v$ ne figure comme variable libre que dans la prémisse auxiliaire (0), les autres prémisses et la conclusion étant des formules quantificationnelles fermées, on peut appliquer les lois de passage et par des tranformations triviales réduire la prémisse auxiliaire à une tautologie dont on se débarrassera grâce à un *modus ponens*.

$$(8) \qquad \exists v((\exists x A x v \supset \forall y \exists x A x y) \wedge (\exists x \forall y A x y)) \supset \forall y \exists x A x y$$
$$(9) \qquad (\exists v(\exists x A x v \supset \forall y \exists x A x y) \wedge (\exists x \forall y A x y)) \supset \forall y \exists x A x y$$
$$(10) \qquad \exists v(\exists x A x v \supset \forall y \exists x A x y) \supset (\exists x \forall y A x y \supset \forall y \exists x A x y)$$
$$(11) \qquad (\forall v \exists x A x v \supset \forall y \exists x A x y) \supset (\exists x \forall y A x y \supset \forall y \exists x A x y)$$
$$(12) \qquad (\forall y \exists x A x y \supset \forall y \exists x A x y) \supset (\exists x \forall y A x y \supset \forall x \exists y A x y)$$
$$(13) \qquad (p \supset p) \supset (\exists x \forall y A x y \supset \forall x \exists y A x y)$$
$$(14) \qquad \text{donc } \exists x \forall y A x y \supset \forall y \exists x A x y$$

Donc la condition (1) suffit pour arriver à la conclusion. On n'a pas besoin de (0).

On aurait pu continuer le processus d'élimination des variables signalisées et se débarasser de $z$ en ajoutant une nouvelle prémisse auxiliaire (0'), qu'on aurait éliminée ensuite de la même manière.

La contrainte interdisant de signaliser deux fois la même variable trouve sa justification dans le fait qu'une infraction à cette règle rend impossible l'élimination de la prémisse auxiliaire.

Considérons la pseudo-démonstration 5.99, dans laquelle la variable d'instantiation $y$ apparaît deux fois.

(5.99)

| | | | |
|---|---|---|---|
| (1) | $\exists x F x$ | | P.A. |
| (2) | $Fy$ | $y$ | 1, I.E. |
| (3) | $\forall z F z$ | $y$ | 1, G.U. |
| (4) | $\exists x F x \supset \forall z F z$ | | Cond(1,3) |

Cette démonstration fait appel aux deux règles suspectes I.E. et G.U. Pour nous débarrasser de ces règles et les remplacer par deux applications du *modus ponens*, nous devons introduire les deux prémisses auxiliaires, notées P.A.1 et P.A.2 :

(0) $\exists x F x \supset Fy$,

et

(0′) $Fy \supset \forall z F z$.

Nous obtenons alors une démonstration correcte (5.100), mais qui prouve autre chose que la formule visée (c'est-à-dire $\exists x F x \supset \forall z F z$).

(5.100)

| | | |
|---|---|---|
| (1) | $Fy \supset \forall z F z$ | P.A. 2 |
| (2) | $\exists x F x \supset Fy$ | P.A. 1 |
| (3) | $\exists x F x$ | P.A. 3 |
| (4) | $Fy$ | M.P. 2,3 |
| (5) | $\forall z F z$ | M.P. 1,4 |
| (6) | $\exists x F x \supset \forall z F z$ | Cond(3,5) |
| (7) | $(\exists x F x \supset Fy) \supset (\exists x F x \supset \forall z F z)$ | Cond(2,6) |
| (8) | $(Fy \supset \forall z F z) \supset ((\exists x F x \supset Fy) \supset (\exists x F x \supset \forall z F z))$ | Cond(1,7) |

La conclusion de la dérivation étant valide, nous pouvons effectuer la suite de transformations (9–11).

(9) $\forall y((Fy \supset \forall z Fz) \supset ((\exists x Fx \supset Fy) \supset (\exists x Fx \supset \forall z Fz)))$
par application de : $\models \Phi$ si et seulement $\models \forall x \Phi$

(10) $\forall y(((Fy \supset \forall z Fz) \wedge (\exists x Fx \supset Fy)) \supset (\exists x Fx \supset \forall z Fz))$
par importation

(11) $\exists y((\underline{Fy \supset \forall z Fz}) \wedge (\underline{\exists x Fx \supset Fy})) \supset (\exists x Fx \supset \forall z Fz)$

A présent, nous sommes arrêtés. Pour transformer les formules soulignées en tautologies, de manière à pouvoir détacher $(\exists x Fx \supset \forall z Fz)$ et prouver ainsi la formule visée, il faudrait, au préalable, soit distribuer $\exists y$ sur $\wedge$, ce qui est interdit :

(5.101) $(\exists \nu(\Phi \wedge \Psi) \not\equiv (\exists \nu \Phi \wedge \exists \nu \Psi))$,

soit restreindre la portée de $\exists y$, ce qui est impossible précisément parce que le *même* "$y$" apparaît à titre d'occurrence libre dans les deux formules soulignées. On voit bien ici que l'interdiction de signaliser deux fois la même variable est une contrainte *nécessaire* à imposer à l'utilisation des règles I.E. et G.U. En effet, si on la transgresse, on se rend incapable de reconstruire la démonstration ou la dérivation de manière à se passer de ces deux règles.

Les contraintes imposées à I.E. et G.U. sont *suffisantes* puisque, lorsqu'elles sont respectées, la démonstration peut être réécrite sans qu'il soit fait usage des règles I.E. et G.U.

### 5.4.7. *Un regard philosophique sur la correction*

Une règle d'inférence est correcte si et seulement si elle ne conduit jamais des prémisses à une conclusion qui ne serait pas une conséquence logique de celles-ci. Dans l'acception la plus étroite de la relation de conséquence logique, cette relation ne peut relier que des propositions. Une proposition $B$ est une conséquence logique d'un ensemble de propositions $\{A_1, \ldots, A_n\}$ si et seulement si toutes les interprétations qui rendent vraies les propositions $A_1, \ldots, A_n$ rendent vraie la proposition $B$. A parler strictement, seules les propositions peuvent être vraies ou fausses. Les fonctions propositionnelles ne le peuvent pas. La relation de conséquence logique mentionnée plus haut — et qui est celle qu'adopte K. Fine — n'est définie que pour les formules fermées. Nous verrons à la section 7.7 comment étendre la relation de conséquence logique aux fonctions propositionnelles.

Les règles G.U. et I.E. de Quine sont correctes au sens défini plus haut : elles ne permettent pas de construire des preuves dans lesquelles on déduirait des formules fermées fausses.

Il s'agit là d'une conception *minimaliste* de la correction et d'une conception *instrumentaliste* de la déduction, que Fine caractérise en ces termes :

> La justification des règles telles que G.U. et I.E. ne dépend pas de la validité des inférences que [ces règles] sanctionnent, mais du fait que leur application ne cause aucun mal (Fine, 1985, p. 126).

On peut avoir des exigences *plus fortes* et adhérer à une conception de la déduction inspirée par un certain idéal de *rationalité*. C'est ce que fait Fine :

> Selon cette conception [celle qu'il défend], la justification de règles telles que G.U. et I.E. repose sur le caractère local des inférences qu'elle sanctionne comme valides, et non sur le caractère du système de règles auquel elles appartiennent (Fine, 1985, *Ibid.*).

Deux conceptions s'affrontent ici : la conception *holiste* de la validité (la validité des parties résulte de la validité du tout), et la conception *atomiste* ou *modulaire* (la validité du tout résulte de la validité des parties). Il est vrai que la première est moins satisfaisante. Elle nous impose de faire des *pas* dans une dérivation qui ne correspondent pas à des relations de dépendance logique. Inversement, elle nous interdit certains pas qui correspondent à des relations de dépendance. En outre, on doit parfois revenir en arrière dans une dérivation pour satisfaire les contraintes sur G.U. et E.I. qui sont des contraintes *globales*. Nous nous bornerons néanmoins à la méthode de déduction naturelle de Quine qui est une méthode holiste parce qu'elle est la plus courante et la plus facile à justifier.

### 5.4.8. *Applications de la méthode de déduction naturelle*

*Application 1*

Reprenons l'exercice 1 résolu par la méthode axiomatique (cf. figure 5.4).

La démonstration en déduction naturelle n'est guère plus facile à trouver que la démonstration par la méthode axiomatique. L'effort d'invention se situe aux lignes (5), (7), (12) et (16), dans le choix des variables d'instantiation. Ce qui guide notre choix des instances de la ligne (7), c'est le souci d'exploiter la prémisse II qui apparaît en (6). L'étape la plus créative est la ligne (5).

Il arrive cependant *souvent* que la preuve en déduction naturelle soit sensiblement plus facile à construire que la preuve axiomatique.

*Application 2*

Reprenons l'application 3 de la méthode axiomatique (cf. figure 5.5).

La démonstration est sensiblement plus facile que par la méthode axiomatique.

*Application 3*

Voici un exercice dû à Manna, dont la résolution est nettement plus facile par la méthode de déduction naturelle que par la méthode axiomatique. Comme la

| | | | |
|---|---|---|---|
| | (1) | $\forall x \forall y(\exists z(Rxz \wedge Ryz) \supset Rxy)$ | Prémisse I |
| | (2) | $\forall x \exists y Rxy$ | Prémisse II |
| | (3) | $\longrightarrow Rab \wedge Rbc$ | P. A. |
| | (4) | $\forall x \forall y \forall z((Rxz \wedge Ryz) \supset Rxy)$ | 1, Echange, loi de passage |
| ** | (5) | $\exists y Rcy$ | I.U., 2 |
| | (6) | $Rcv$ | I.E., 5, $v$ |
| * | (7) | $(Rcv \wedge Rcv) \supset Rcc$ | I.U. 4 (×3) |
| | (8) | $Rcv \wedge Rcv$ | Idemp. 6 |
| | (9) | $Rcc$ | M.P. 7, 8 |
| | (10) | $Rbc$ | Simp. 3 |
| | (11) | $Rcc \wedge Rbc$ | Conj. 9, 10 |
| * | (12) | $(Rcc \wedge Rbc) \supset Rcb$ | I.U. 4 (×3) |
| | (13) | $Rcb$ | M.P. 11, 12 |
| | (14) | $Rab$ | Simp. 3 |
| | (15) | $Rab \wedge Rcb$ | Conj. 13, 14 |
| * | (16) | $(Rab \wedge Rcb) \supset Rac$ | I.U. 4 (×3) |
| | (17) | $Rac$ | M.P. 15, 16 |
| | (18) | $(Rab \wedge Rbc) \supset Rac$ | Décharge 3 |

**Figure 5.4** *Exercice 1 résolu par la méthode axiomatique*

| | | | |
|---|---|---|---|
| | (1) | $\longrightarrow \forall x \forall y(Rxy \supset \neg Ryx)$ | P.A. |
| | (2) | $\forall y(Rxy \supset \neg Ryx)$ | I.U., 1 |
| | (3) | $Rxx \supset \neg Rxx$ | I.U., 2 |
| | (4) | $\neg Rxx \vee \neg Rxx$ | Impl. mat. |
| | (5) | $\neg Rxx$ | Idemp. |
| | (6) | $\forall x \neg Rxx$ | G.U. 5, $x$ |
| | (7) | $\forall x \forall y(Rxy \supset \neg Ryx) \supset \forall x \neg Rxx$ | Elimination de P.A. |

**Figure 5.5** *Application 3 de la méthode axiomatique*

déduction naturelle n'est pas linéaire mais bidimensionnelle, on peut construire des dilemmes emboîtés dans d'autres dilemmes et des preuves par l'absurde *emboîtées* dans des preuves conditionnelles, sans jamais perdre de vue le *niveau* auquel on se trouve.

Manna (1974, p. 160) donne 7 prémisses :

(1) $\forall x(x = x)$
(2) $\forall x \forall y \forall z((x = y \wedge y = z) \supset x = z)$
(3) $\forall x \forall y(x = y \supset y = x)$
(4) $\forall x \forall y \forall z((x < y \wedge y < z) \supset x < z)$
(5) $\forall x \forall y(x < y \vee x = y \vee y < x)$
(6) $\forall x \forall y(x = y \supset x \not< y)$
(7) $\forall x \forall y(x < y \supset y \not< x)$;

de ces prémisses, il nous propose de déduire

$$\forall x \forall y \forall z(((x < y \vee x = y) \wedge (y < z)) \supset (x < z)).$$

Une démonstration en déduction naturelle est présentée à la figure 5.6.

*Application 4*

Il s'agit de démontrer que $\forall$ se distribue sur $\wedge$ :

$$\vdash \forall x(Fx \wedge Gx) \equiv \forall xFx \wedge \forall xGx.$$

| | | |
|---|---|---|
| (1) | $\forall x(Fx \wedge Gx)$ | P.A. |
| (2) | $Fx \wedge Gx$ | I.U., 1 |
| (3) | $Fx$ | Simp., 2 |
| (4) | $\forall xFx$ | G.U. 3, $x$ |
| (5) | $Fy \wedge Gy$ | I.U., 1 |
| (6) | $Gy$ | Simp. 5 |
| (7) | $\forall xGx$ | G.U. 5, $y$ |
| (8) | $\forall xFx \wedge \forall xGx$ | Conj. 4, 7 |
| (9) | $\forall x(Fx \wedge Gx) \supset \forall xFx \wedge \forall xGx$ | Déch. |
| (10) | $\forall xFx \wedge \forall xGx$ | P.A. |
| (11) | $\forall xFx, \forall xGx$ | Simp., 10 |
| (12) | $Fx$ | I.U., 11 |
| (13) | $Gx$ | I.U., 11 |
| (14) | $Fx \wedge Gx$ | Conj. 12, 13 |
| (15) | $\forall x(Fx \wedge Gx)$ | G.U. 14, $x$ |
| (16) | $(\forall xFx \wedge Gx) \supset \forall x(Fx \wedge Gx)$ | |
| (17) | $9 \wedge 16$ | Conj. 9, 16 |
| (18) | $\forall x(Fx \wedge Gx) \equiv (\forall xFx \wedge \forall xGx)$ | Equiv. mat. 17 |

*Remarque.* Les lignes (4) et (15) contiennent la même variable signalisée $x$. La démonstration n'est cependant pas erronée. Ce qui est interdit, c'est d'utiliser plusieurs fois une variable signalisée au sein d'une *même dérivation achevée*. Or cette interdiction est respectée : les dérivations (1-8) et (10-15) sont, en effet,

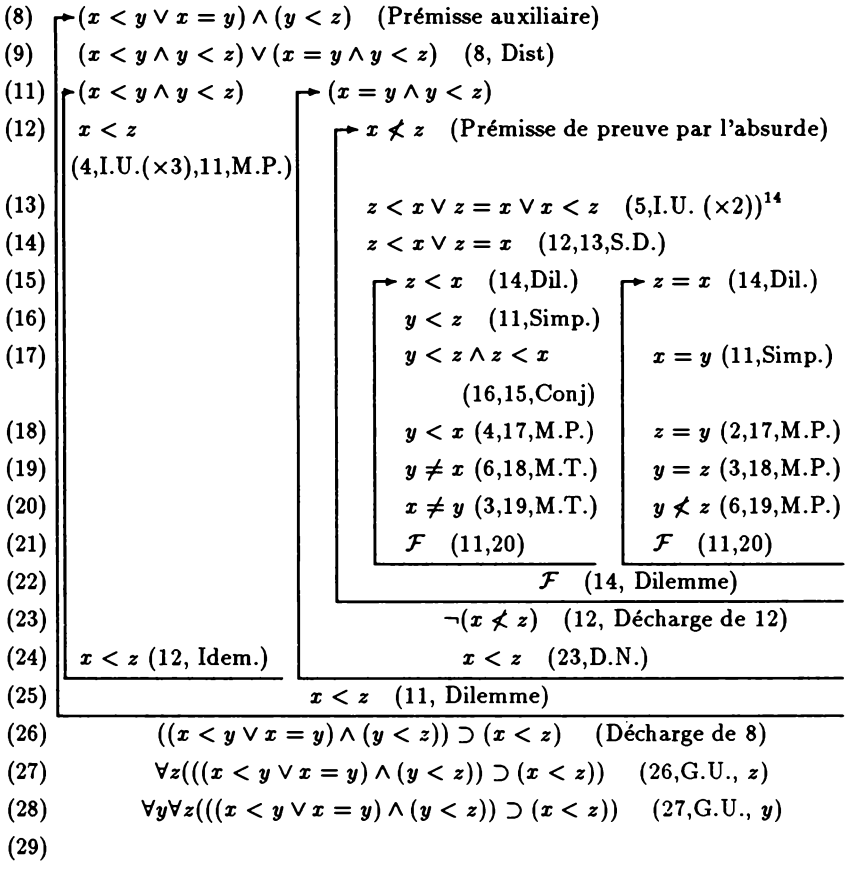



**Figure 5.6** *Démonstration en déduction naturelle*

[14]Nous n'indiquerons plus les applications de I.U. dans les lignes suivantes.

des dérivations ***différentes***. Par ailleurs, dans la première de ces dérivations, on a fait deux applications de G.U., mais en prenant soin d'utiliser des variables d'instantiation différentes (respectivement $x$ et $y$). C'est dans ce but que nous avons instancié deux fois les prémisses auxiliaires.

*Application 5*

Exercice : remplir les blancs et signaliser.

| | | |
|---|---|---|
| (1) | $Fxz$ | P.A. |
| (2) | $\cdots$ | G.U. |
| (3) | $\cdots$ | G.U. |
| (4) | $\cdots$ | I.U. |
| (5) | $\cdots$ | I.U. |
| (6) | $Fxz \supset Fzy$ | |
| (7) | $\exists z(Fxz \supset Fzy)$ | |

# Chapitre 6
# Le calcul des prédicats – méthodes analytiques –

Au début du chapitre 4, nous avons mis en évidence l'importance de la *normalisation* par laquelle une dérivation peut être mise en forme "normale" qu'il s'agisse d'un arbre de séquents *sans coupure* ou d'une dérivation en déduction naturelle *sans détour*. Nous avions souligné à cette occasion que les résultats de Prawitz avaient comblé le vide qui existait initialement entre les *méthodes synthétiques* de démonstration dans lesquelles on construit la démonstration d'un théorème à partir d'éléments exogènes (axiomes, théorèmes, prémisses auxiliaires) et les *méthodes analytiques* dans lesquelles la démonstration est faite en démembrant la négation du théorème à démontrer en ses parties puis en reconstituant une preuve directe par "remembrement".

L'existence d'un théorème de normalisation pour la déduction naturelle qui est le pendant du théorème d'élimination de la coupure pour le calcul des séquents permet de dire que, vues de l'intérieur, les *méthodes synthétiques* et les *méthodes analytiques*, pour ce qui est du calcul des propositions et du calcul des prédicats, ont en commun un trait essentiel : elles sont *normalisables*.

Si l'on ajoute aux axiomes et règles du calcul des prédicats des axiomes extra-logiques tels que les axiomes de l'arithmétique, parmi lesquels figure le schéma d'induction mathématique $((\Phi_0 \wedge \forall x(\Phi x \supset \Phi S(x))) \supset \forall x \Phi x)$, la normalisation n'est plus possible que dans certains cas et sous certaines conditions. L'adjonction à la logique des prédicats de ce "corps étranger" que constitue l'induction mathématique agit comme un prisme qui différencie les méthodes synthétiques des méthodes analytiques (qui cessent d'être entièrement analytiques puisque la coupure n'est plus éliminable entièrement et inconditionnellement).

Si l'adjonction d'un principe mathématique aux règles de logique retroagit

de la sorte sur celles-ci, on doit réviser la thèse reçue selon laquelle il y a une cloison étanche entre la partie spécifiquement mathématique d'une théorie (les axiomes et les théorèmes) et les liens de dépendance logique entre théorèmes et axiomes. Nous reviendrons sur ce point à la section 8.8.

## 6.1. La méthode des tableaux sémantiques

Nous avons introduit en 3.3.3 la notion de *constituant* ou *sous-formule*. En 4.1.15, nous avons exposé les règles qui, en démembrant les formules composées, font remonter à la surface les contradictions sous la forme de littéraux complémentaires. Nous n'avons eu recours qu'à deux sortes de règles : les règles de *prolongation* (dites de type $\alpha$) et les règles de *ramification* (dites de type $\beta$).

### 6.1.1. *Constituants syntaxiques et constituants logiques*

La méthode des tableaux sémantiques pour le calcul des propositions consistait uniquement en règles de *décomposition* des formules. L'application répétée de ces règles à des formules moléculaires de longueur finie ou à des ensembles finis de telles formules permettait soit de prouver qu'*aucune* assignation de valeur de vérité aux constituants atomiques de ces formules n'existait, soit qu'*une* telle assignation *au moins* existait.

Dans le calcul des propositions, toute formule bien formée est de longueur *finie* et les constituants *logiques* d'une expression moléculaire coïncident avec les constituants *syntaxiques*, c'est-à-dire avec les *parties* de la formule que l'on peut rattacher à un noeud dans un arbre syntaxique.

Sur ces deux points, l'extension de la méthode des tableaux sémantiques au calcul des prédicats apporte des changements. D'une part, les règles de décomposition des propositions générales ne seront plus vraiment des règles de décomposition, mais seront plutôt des règles d'exemplification. D'autre part, les notions de "constituants logiques" et de "parties" ou "constituants syntaxiques" ne coïncident plus :

> les *parties* immédiates de $\forall x A$ sont $\forall x$ et $A$, mais les *constituants* [au sens logique] de $\forall x A$ sont les différentes propositions $A(k)$, où $k$ nomme un membre dans [l'univers] $U$ (Robinson, 1969, p. 19).

Par exemple, si $U$ est l'ensemble des entiers naturels, $A(0), A(1), A(2), \ldots$ sont des constituants logiques de $\forall x A$. Ces deux différences sont d'ailleurs liées. C'est parce que $A(0), A(1), \ldots$ ne sont pas des parties syntaxiques de $\forall x A$ que la règle d'exemplification n'est pas une règle de décomposition au sens strict.

### 6.1.2. *Les règles de décomposition d'une universelle*

Sachant qu'une universelle $\forall x A$ est équivalente à une conjonction contenant autant de propositions atomiques qu'il y a de membres $k$ dans l'*univers* $\mathcal{U}$, la règle de *décomposition* d'une universelle sera une sorte de règle de *prolongation* comme celle qui gouverne le $\wedge$, mais de prolongation qui peut continuer *indéfiniment* : l'*arbre* auquel donne lieu la décomposition d'une formule isolée ou d'un ensemble *fini* de formules pourra être un *arbre infini*.

Dans la figure 4.8 de 4.1.16, nous avions abrégé $\|\phi \circ \psi\| = \mathbf{V}$ en $\mathbf{V}(\phi \circ \psi)$ pour tout connecteur $\circ$. Nous ferons de même ici. Nous abrégerons $\|Qx A\| = \mathbf{V}$ en $\mathbf{V}Qx A$ pour tout quantificateur $Qx$ (et de même pour $\mathbf{F}$).

Pour laisser ouverte la possibilité de *prolonger* de manière répétée $\mathbf{V}\forall x A$ en

$$\begin{array}{l} \mathbf{V}A(k_0/x), \\ \mathbf{V}A(k_1/x), \\ \vdots \end{array}$$

on s'interdira de *cocher* la ligne générique $\mathbf{V}\forall x A$.

En logique bivalente classique, $\mathbf{F}\exists x A$ est équivalent à $\mathbf{V}\neg\exists x A$, et comme

(6.1) $\quad \neg\exists x A \equiv \forall x \neg A,$

$\mathbf{F}\exists x A$ est, en fait, une universelle. On pourra donc l'instancier indéfiniment :

(6.2) $\quad \begin{array}{l} \mathbf{F}\exists x A, \\ \mathbf{F}A(k_0/x), \\ \mathbf{F}A(k_1/x), \\ \vdots \end{array}$

et on retrouvera également l'interdiction de *cocher* la ligne générique $\mathbf{F}\exists x A$.

Pour marquer la ressemblance, nous placerons ces deux règles dans une même rubrique : ce sont des règles d'*instantiation* d'une *universelle* (quasi-universelle dans le deuxième cas) appelées règles du *type* $\gamma$, schématisées comme suit :

(6.3) $\quad \dfrac{\gamma}{\gamma_0(a)}$ , où $a$ est n'importe quel paramètre.

Exploitant la convention selon laquelle les évaluations $\mathbf{V}X$ (respectivement $\mathbf{F}X$) s'expriment en plaçant $X$ à gauche (respectivement à droite) de la branche de l'arbre sur laquelle se trouve le noeud étiqueté par $X$, nous représenterons les deux règles $\gamma$ à l'aide du graphique de la figure 6.1.

### 6.1.3. *La notion d'exemplification*

Toute proposition générale possède, en tant que proposition, une *valeur de vérité*, et en tant que générale, une ou plusieurs *instances*. L'instance (ou les

$$\frac{\forall x A}{Aa} \qquad \frac{\exists x A}{Aa}$$

**Figure 6.1** *Les règles de type* $\gamma$

instances) d'une proposition générale n'a pas nécessairement la même valeur de vérité que celle-ci. Par exemple, sachant que l'univers de $x$ est $\mathcal{D}$, où $\mathcal{D}$ est l'ensemble des naturels, "4 est pair" et "5 est pair" sont deux instances de la proposition générale *fausse* "Tout est pair", mais seule la deuxième a, par surcroît, la même valeur de vérité que la proposition instanciée. Robinson (1979, p. 26) donne le nom d'*instance exemplifiante* aux propositions atomiques qui entretiennent avec une proposition générale les *deux* relations distinguées plus haut : instancier et avoir en partage la même valeur de vérité.

Toute proposition générale[1], qu'elle soit vraie ou fausse, a au moins une instance exemplifiante dans laquelle la variable quantifiée est remplacée par un terme singulier $a$. La logique classique exclut l'univers vide.

### 6.1.4. *Décomposition d'une particulière*

Quand la proposition à exemplifier est une particulière $\exists x A$, il se peut qu'il n'existe qu'une seule instance exemplifiante — c'est le cas pour $\exists x A$, où $A$ est la fonction propositionnelle $x + x = x$ et $\mathcal{D}$ est $\Re$ — aussi la règle de décomposition sera-t-elle une règle assujettie à la même contrainte que la règle I.E. de la déduction naturelle. Les schémas d'inférence

(6.4) $\mathbf{V}\exists x A$
$\mathbf{V}Aa$ ($a$ est inédit),

(6.5) $\mathbf{F}\forall x A$
$\mathbf{F}Aa$ ($a$ est inédit),

ne pourront s'appliquer qu'une fois, et le terme singulier $a$ devra ne pas avoir figuré antérieurement sur la branche. On *doit* cocher la ligne générique dès qu'on a généré une instance.

Les règles d'*instantiation* d'une *existentielle* (ou quasi-existentielle) sont rangées sous le *type* $\delta$ et schématisées comme suit :

(6.6) $\dfrac{\delta}{\delta_0(a)}$, où $a$ est un paramètre inédit.

Dans la notation graphique que nous avons adoptée les deux règles $\delta$ se représentent par la figure 6.2.

[1] Les propositions *générales* se subdivisent en propositions *universelles* (de la forme $\forall x A$) et *particulières* (de la forme $\exists x A$).

$$\frac{\exists x A}{Aa} \qquad \frac{\forall x A}{Aa}$$

**Figure 6.2** *Les règles de type δ*

### 6.1.5. *Tableau récapitulatif des règles*

Ce tableau est présenté à la figure 6.3.

| décomposition de propositions moléculaires | | décomposition de propositions générales | |
|---|---|---|---|
| *Règles α* | *Règles β* | *Règles γ* | *Règles δ* |
| prolongation | ramification | instantiation d'universelle ou de quasi-univ. (I.U.) | instantiation d'existentielle ou de quasi-exist. (I.E.) |
| $\dfrac{\alpha\surd}{\begin{array}{l}\alpha_1\\ \alpha_2\end{array}}$ | $\dfrac{\beta\surd}{\beta_1 \quad \beta_2}$ | $\dfrac{\gamma}{\gamma_0(a)}$ | $\dfrac{\delta\surd}{\delta_0(a)}$ |
| | | $a$ est un paramètre quelconque | $a$ est un paramètre inédit |

**Figure 6.3** *Règles* $\alpha, \beta, \gamma, \delta$

### 6.1.6. *Propositions passives et propositions actives*

Appelons *degré de complexité* d'une formule le *nombre*[2] de ses constituants. L'élimination d'un connecteur ou d'un quantificateur entraîne la disparition

[2] On ne fait donc pas intervenir ici le nombre d'opérateurs emboîtés comme c'est le cas dans la définition de la *profondeur* (cf. 3.4.9)

d'un et d'un seul constituant. Toutes ces règles font donc passer une formule de complexité $n$ à une formule de complexité $n-1$. Il ne s'ensuit pas, cependant, que l'application répétée de ces règles à une formule — qui par définition a une longueur finie —, ou à un ensemble fini de formules, génère toujours un *arbre fini*. C'était le cas dans le calcul des propositions, où l'on ne disposait que de règles $\alpha$ et $\beta$ qui ne peuvent être appliquées qu'une seule fois à la même formule. Ce n'est plus le cas dans le calcul des prédicats, où les règles de type $\gamma$ peuvent être réappliquées indéfiniment à la même ligne générique. C'est pourquoi on ne coche pas la ligne générique que l'on a instantiée, en appliquant une règle de type $\gamma$. Les règles de type $\gamma$ peuvent générer *indéfiniment* de nouveaux constituants qui ne sont pas des *parties* ou *constituants syntaxiques* de la formules générique. Pour souligner cette particularité, Robinson appelle *propositions actives* les lignes génériques $\mathbf{V}\forall xA$ et $\mathbf{F}\exists xA$ qui servent d'*entrée* aux règles de type $\gamma$.

Les règles $\alpha$, $\beta$ et $\delta$ ont ceci en commun : on ne peut les appliquer qu'une seule fois à une formule donnée, elles ne peuvent rien *créer*. Les règles $\alpha$ et $\beta$ *démembrent* une ligne par l'élimination d'un connecteur. Les règles $\delta$ réécrivent la ligne sous la forme d'une instance après élimination du quantificateur. Pour souligner ce qu'il y a de commun aux propositions de type $\alpha$, $\beta$, et $\delta$, Robinson les appelle *propositions passives*.

La présence de règles de type $\gamma$ fait que, même si l'entrée est une formule *unique* — donc une expression de longueur finie — ou un ensemble *fini* de formules, il peut exister des arbres infinis (cf. figure 6.4).

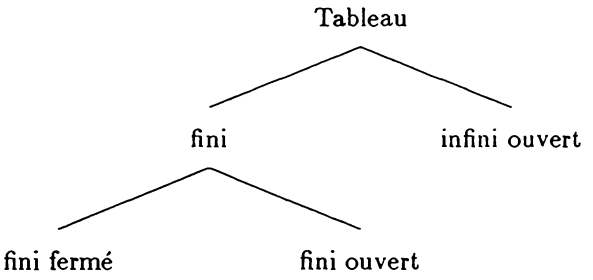



**Figure 6.4** *Les trois sortes de tableaux*

La figure 6.5 présente un exemple de tableau infini ouvert.

Remarquons que la règle $\delta$ n'est jamais appliquée qu'*une seule fois* à une *même* formule sur la même branche (les formules $\exists y\Phi ay$, $\exists y\Phi by$, $\exists y\Phi cy$ sont des formules *différentes*).

$\forall x \exists y \Phi xy$
$\exists y \Phi ay \surd$
$\Phi ab$
$\exists y \Phi by \surd$
$\Phi bc$
$\exists y \Phi cy \surd$
$\Phi cd$
$\vdots$

**Figure 6.5** *Tableau infini ouvert*

### 6.1.7. *Nécessité d'un plan de travail*

Considérons l'ensemble suivant de propositions :

(6.7) $\{\forall x \neg Fx, \exists x Fx\}$.

Cet ensemble est sémantiquement inconsistant, mais si nous appliquons les règles $\gamma$ et $\delta$ d'une manière anarchique, nous pouvons tirer une infinité d'instances de ces deux formules qui ne seront jamais des littéraux complémentaires. C'est le cas dans la figure 6.6.

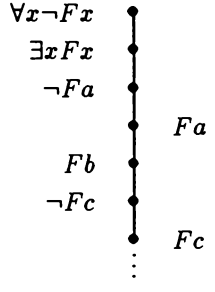



**Figure 6.6**

Pour arriver à une *contradiction flagrante*, c'est-à-dire à deux littéraux complémentaires, il faut que la variable liée des deux formules $\forall x \neg Fx$ et $\exists x Fx$ soit instanciée par le *même* paramètre. Il faudra donc que le paramètre introduit par l'instantiation de ces formules soit *réutilisé* par l'autre. Or le paramètre introduit par l'application d'une règle $\delta$ doit être inédit.

Pour faire apparaître une contradiction flagrante dans le cas qu'on vient d'évoquer, il faudra :

a) appliquer la règle $\delta$ *avant* la règle $\gamma$,

b) utiliser comme paramètre d'instantiation pour $\gamma$ non pas n'importe quel paramètre, mais un paramètre *déjà introduit* par $\delta$, comme on l'a fait

dans l'arbre de la figure 6.7.

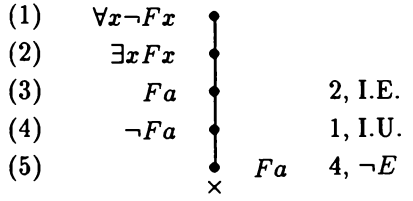



**Figure 6.7**

Faute d'adopter un plan de travail, on risque d'engendrer des arbres infinis à partir d'un ensemble de formules dont on aurait pu tirer un *arbre fini fermé* grâce à un choix plus avisé. On risque également d'engendrer des arbres infinis à partir d'un ensemble de formules dont on aurait pu tirer un arbre *fini ouvert* si on avait adopté un bon plan de travail.

Considérons l'ensemble suivant de propositions :

(6.8) $\{\forall x \neg Fx, \exists x Gx\}$.

On pourrait, par un choix avisé, construire un modèle muni d'un domaine fini dans lequel il est satisfait. Une application anarchique des règles d'instantiation donnera l'arbre infini de la figure 6.8.

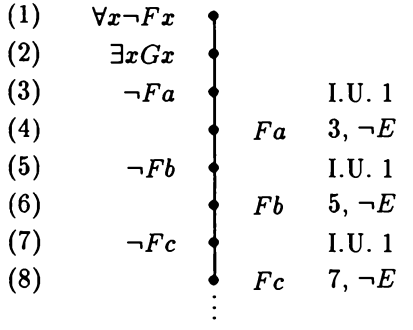



**Figure 6.8**

Si, au contraire, nous commençons par appliquer la règle $\delta$ *et* que nous puisons les paramètres pour les propositions $\gamma$ dans l'ensemble des paramètres générés par les règles $\delta$, nous obtenons l'arbre *fini ouvert* de la figure 6.9.

On voit que l'ensemble $\{\forall x \neg Fx, \exists x Gx\}$ est sémantiquement consistant : il est satisfait pour les évaluations $||Ga|| = \mathbf{V}$ et $||Fa|| = \mathbf{F}$.

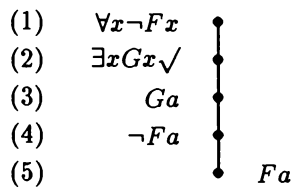



**Figure 6.9**

Le *plan de travail* se résume dans les maximes suivantes : "priorité aux formules passives" c'est-à-dire aux formules $\alpha, \beta, \delta$[3], et "réutilisez pour l'instantiation des formules $\gamma$ les paramètres introduits par l'instantiation des formules $\delta$ mais n'utilisez pas deux fois sur une même branche le même paramètre pour l'instantiation d'une proposition active (formule $\gamma$)".

### 6.1.8. *L'univers auto-généré*

Les paramètres qui serviront à instancier les propositions $\gamma$ sont les paramètres inédits introduits lors de chaque nouvelle application d'une règle $\delta$. Le lexique s'accroît ainsi progressivement. En ce sens, le lexique est "auto-généré". S'il n'existe pas de proposition $\delta$ capable de générer un paramètre et qu'il y a des propositions $\gamma$, on est bien forcé, pour initialiser le processus, d'instancier les propositions $\gamma$ au moyen d'un paramètre inédit $a$.

Même si l'on se borne à n'utiliser comme paramètre d'instantiation des formules $\gamma$ que les paramètres introduits par l'application des formules $\delta$, on aura dans certains cas un *lexique infini*. Par exemple, la formule

(6.9) $\forall x \exists y Fxy$

génère inévitablement, lors de l'instantiation, une branche infinie et un lexique d'instantiation infini, comme on le voit dans la figure 6.10.

### 6.1.9. *L'algorithme de construction de tableaux*

L'algorithme[4] de la figure 6.11 repose sur le précepte "Travaillez d'abord sur les propositions passives".

*Remarques.*

[3] Ces formules sont traitées par des règles qui ne créent rien, mais qui se bornent à *démembrer*.

[4] Adapté de H. Leblanc et de W. Wisdom, 1972, p. 189.

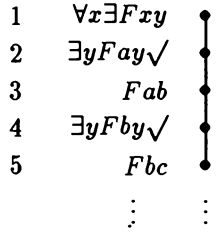



**Figure 6.10**

- L'algorithme admet trois types de comportement : arrêt en I, arrêt en II, pas d'arrêt.
- On peut perfectionner l'algorithme (cf. Rivenc, 1989, p. 243) en décomposant l'instruction-test figurant dans le deuxième losange en deux instructions-tests : l'une pour les formules $\alpha$ et $\delta$, l'autre pour les formules $\beta$. De cette manière, les prolongations précèdent toujours les ramifications. Ceci écourte les arbres (voir les modules de la figure 6.12).

### 6.1.10. *Exemple d'arbre fini fermé*

Démontrons que l'ensemble des trois propositions 6.10 est sémantiquement inconsistant :

(6.10) (1) $\forall x \forall y \forall z((\text{Père}(x,y) \wedge \text{Père}(y,z)) \supset \text{Grand-Père}(x,z))$,
(2) $\forall x \exists y \text{Père}(y,x)$,
(3) $\neg \forall x \exists y \text{Grand-Père}(y,x)$.

Dans les colonnes de droite de la figure 6.13 figurent :

(1) le type de règle $(\alpha, \beta, \gamma, \delta)$,
(2) le nom de la ligne-entrée,
(3) le paramètre.

*Remarque.* A la ligne (13), nous avons sauté toutes les applications inutiles de la règle $\gamma$ à la ligne (1) à partir du lexique $\{a, b, c\}$. L'application mécanique de l'algorithme aurait conduit à un arbre beaucoup plus long, mais chaque branche aurait été néanmoins flanquée, à sa gauche et à sa droite, d'un même atome.

### 6.1.11. *Exemple d'arbre fini ouvert*

Le raisonnement 6.11 est non valide. L'ensemble formé par ses prémisses et la négation de la conclusion est sémantiquement consistant, comme le montre

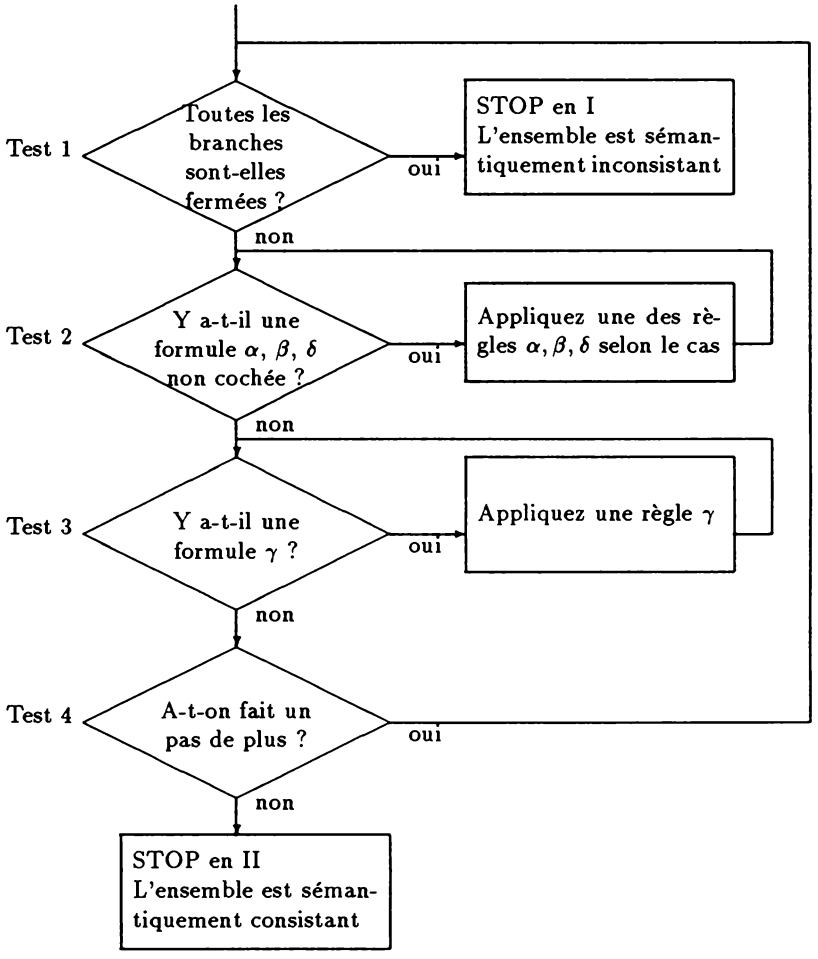



**Figure 6.11** *Plan de travail*

l'arbre de la figure 6.14. La branche ouverte de gauche montre que l'existence d'un objet $a$ qui est à la fois vertébré et non-homme est compatible avec les prémisses du raisonnement 6.14, contrairement à ce qu'affirme la conclusion de ce raisonnement. La branche ouverte de droite montre que l'existence d'un objet $a$ qui est à la fois non-homme, mammifère et vertébré est compatible avec ces mêmes prémisses.

(6.11) Tous les hommes sont des mammifères,
Tous les mammifères sont des vertébrés,
donc tous les vertébrés sont des hommes.

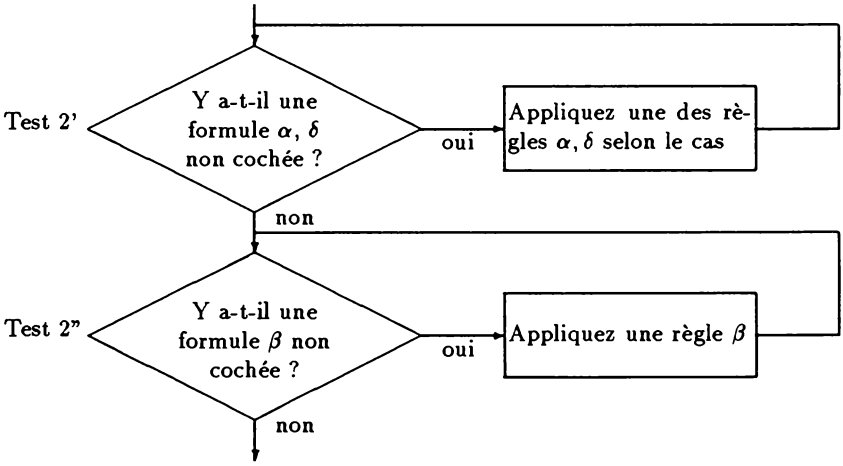



**Figure 6.12** *Modules*

### 6.1.12. *Exemple d'arbre infini ouvert*

La formule 6.12 n'est pas valide. Sa négation est sémantiquement consistante, comme le montre l'arbre de la figure 6.15.

(6.12) $\forall x \exists y(x < y) \supset \exists y \forall x(x < y)$.

On voit que le lexique d'instantiation continuera à s'accroître indéfiniment tant à gauche de la branche qu'à droite. L'arbre est infini.

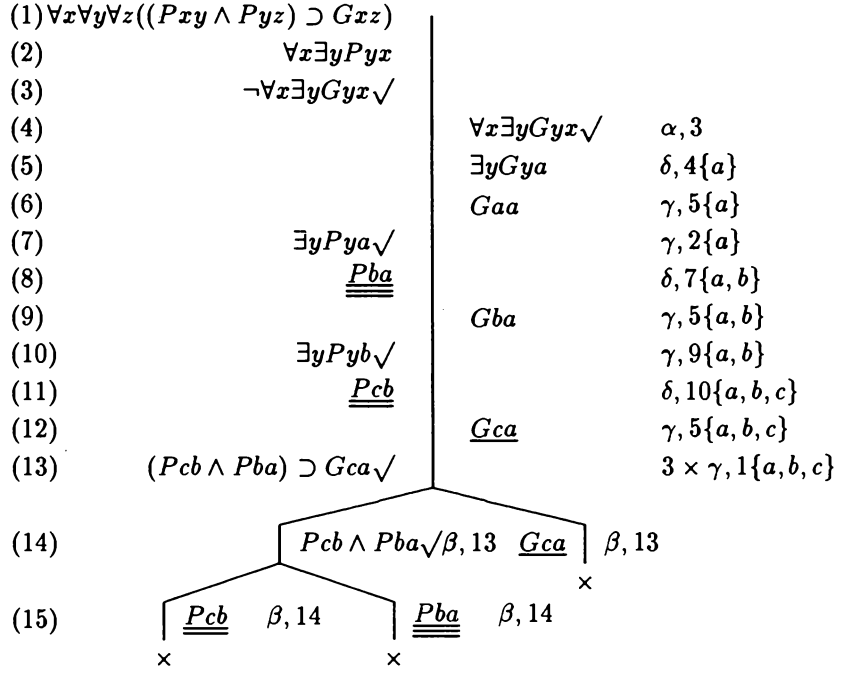
## 6.2. La méthode des séquents

Les notions-clés de la méthode des tableaux sémantiques pour le calcul des prédicats interviennent dans l'exposé, plus complexe, de la méthode des séquents, méthode qui est historiquement antérieure à celle des tableaux. Rappelons que la méthode des séquents permet de construire des *démonstrations* axiomatiques *directes* et pas uniquement des *réfutations*, comme la méthode des tableaux.

### 6.2.1. *Les règles gouvernant les quantificateurs*

Aux règles d'introduction de *connecteurs* à gauche et à droite de la flèche des séquents (cf. section 4.1.2) viennent s'ajouter, dans le calcul des prédicats, les quatre règles d'introduction des quantificateurs de la figure 6.16.

On se souvient que les règles du calcul des séquents sont réversibles. Lues de bas en haut, ces règles apparaissent comme des règles d'*élimination*. Les




*Remarque.* : on notera que le lexique d'instantiation s'accroît progressivement : {a}, {a, b}, {a, b, c}. Il est sous-entendu dans les lignes 14 et 15 pour alléger la notation.

**Figure 6.13** *Arbre fini fermé*

règles (2) et (3) sont des règles d'*instantiation d'une existentielle* (ou quasi-existentielle), c'est-à-dire, pour reprendre la terminologie de la méthode des tableaux, des règles de type $\delta$ où $b$ doit être un terme *inédit*, contrairement au $r$ des règles (1) et (4) de la figure 6.16. Les règles (1) et (4) sont des règles d'*instantiation d'une universelle* (ou quasi-universelle), c'est-à-dire des règles de type $\gamma$. On se souvient que les formules servant d'entrée aux règles de type $\alpha$ $\gamma$ ne pouvaient pas être *cochées*. En effet, elles sont réutilisables. Dans le calcul des séquents, cette latitude que l'on a de *réinstancier* $\forall xA, \alpha \Longrightarrow \beta$ ou $\alpha \Longrightarrow \beta, \exists xA$ s'exprime par le fait que la formule générique est *répétée* dans la ligne de la formule instantiée. Cette répétition du quantificateur est l'équivalent de l'*interdiction de cocher* les formules de type $\gamma$ dans la méthode des tableaux.

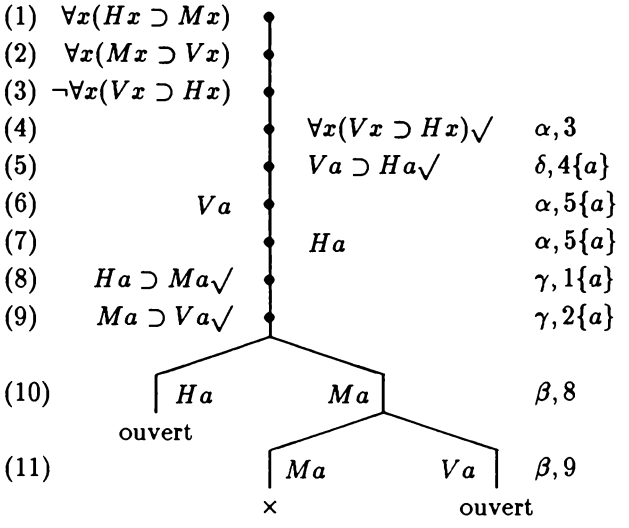



**Figure 6.14** *Arbre fini ouvert*

### 6.2.2. *Plan de travail d'une preuve par séquents*

Dans la construction d'une preuve par séquents, l'arbre est orienté comme dans la nature : la racine est en bas et les feuilles en haut. A la racine de l'arbre, on écrit la formule ou l'ensemble de formules dont on cherche à établir l'inconsistance, et l'on "fait grandir" l'arbre en éliminant les connecteurs et les quantificateurs à l'aide des règles d'inférence des séquents lues de bas en haut. On applique les mêmes maximes que dans la méthode des tableaux (priorité aux formules passives, instantiation des formules $\gamma$ à l'aide de tous les paramètres introduits par les formules $\delta$ précédentes, sauf s'il y a déjà une instance $\gamma$ sur la branche sur laquelle on travaille).

Kleene a introduit un perfectionnement dans le plan de travail, le groupement en *rounds* ou *reprises* :

> Les pas le long d'un chemin particulier seront groupés en reprises [...] En effectuant la reprise $d$, nous prenons tour à tour chacune des occurrences de formule $A_1 \ldots A_m$, $B_1 \ldots B_n$ et [...] nous faisons si c'est possible un pas ou un nombre fini de pas avec cette formule comme formule principale (à moins qu'en cours de route ce chemin ne se ferme [...]) (Kleene 1967, 1972, pp. 303-304)

L'idée est la suivante : on commence par appliquer *une fois* une règle $\alpha$ (respectivement $\beta$ et $\delta$) aux formules de type $\alpha$ (respectivement $\beta$ et $\delta$), c'est-à-dire aux formules passives. Ce pas accompli sur toutes les formules $\alpha$, $\beta$, $\delta$ disponibles, on s'occupe des formules actives ($\gamma$). On ne traite donc pas les

| | | | |
|---|---|---|---|
| (1) | | $\forall x \exists y(x < y) \supset \exists y \forall x(x < y)\surd$ | |
| (2) | $\forall x \exists y(x < y)$ | | $\alpha, 1$ |
| (3) | | $\exists y \forall x(x < y)$ | $\alpha, 1$ |
| (4) | $\exists y(a < y)$ | | $\gamma, 2\{a\}$ |
| (5) | $a < b$ | | $\delta, 4\{a, b\}$ |
| (6) | | $\forall x(x < a)$ | $\gamma, 3\{a, b\}$ |
| (7) | | $\forall x(x < b)$ | $\gamma, 3\{a, b\}$ |
| (8) | | $c < a$ | $\delta, 6\{a, b, c\}$ |
| (9) | | $c < b$ | $\delta, 7\{a, b, c\}$ |
| (10) | | $d < a$ | $\delta, 6\{a, b, c, d\}$ |
| (11) | | $d < b$ | $\delta, 7\{a, b, c, d\}$ |
| (12) | $\exists y(b < y)$ | | $\gamma, 2\{a, b, c, d\}$ |
| (13) | $\exists y(c < y)$ | | $\gamma, 2\{a, b, c, d\}$ |
| (14) | $\exists y(d < y)$ | | $\gamma, 2\{a, b, c, d\}$ |
| (15) | $b < e$ | | $\delta, 10\{a, b, c, d, e\}$ |
| (16) | $c < e$ | | $\delta, 11\{a, b, c, d, e\}$ |
| (17) | $d < e$ | | $\delta, 12\{a, b, c, d, e\}$ |
| (18) | | $\forall x(x < c)$ | $\gamma, 3\{a, b, c, d, e\}$ |
| (19) | | $\forall x(x < d)$ | $\gamma, 3\{a, b, c, d, e\}$ |
| (20) | | $\forall x(x < e)$ | $\gamma, 3\{a, b, c, d, e\}$ |

**Figure 6.15** *Arbre infini ouvert*

$$(1)\ \forall \Longrightarrow \frac{A(r), \forall x A, \alpha \Longrightarrow \beta}{\forall x A, \alpha \Longrightarrow \beta} \qquad (3)\ \frac{\alpha \Longrightarrow \beta, A(b)}{\alpha \Longrightarrow \beta, \forall x A} \Longrightarrow \forall$$

$$(2)\ \exists \Longrightarrow \frac{A(b), \alpha \Longrightarrow \beta}{\exists x A, \alpha \Longrightarrow \beta} \qquad (4)\ \frac{\alpha \Longrightarrow \beta, \exists x A, A(r)}{\alpha \Longrightarrow \beta, \exists x A} \Longrightarrow \exists$$

**Figure 6.16** *Règles des quantificateurs pour les séquents*

nouvelles formules passives qui pourraient être issues de l'application des règles $\alpha, \beta$ et $\delta$. On s'occupe, au contraire, des formules actives qui restent et on fait autant d'instantiations avec celles-ci qu'il y a de paramètres dans le lexique. A tout moment, il n'y a qu'un nombre fini de paramètres dans le lexique : ce sont soit les paramètres introduits par des applications de la règle $\delta$, soit des

variables libres ou des constantes d'individus qui figuraient au départ dans les formules $A_1 \ldots A_m$, $B_1 \ldots B_n$. Or ces variables et ces constantes sont en nombre fini car les formules en question sont en nombre fini. Nous verrons cependant plus loin que la méthode des séquents peut être étendue aux ensembles infinis.

Un chemin se ferme quand on arrive à un axiome, c'est-à-dire à un séquent dans lequel il y a un même atome à gauche et à droite de la branche, plus, éventuellement, d'autres formules non atomiques.

*Exemple.* Démontrons la validité de l'énoncé

(6.13) S'il y a un philosophe que tous les philosophes contredisent, alors il y a un philosophe qui se contredit lui-même (Quine, 1973, p. 150)

par la méthode des séquents. L'arbre de la figure 6.17 est une preuve par séquents de l'énoncé 6.13. Nous représenterons "Philosophe" par le symbole de prédicat à une place $F(-)$ et "contredit" par le symbole de prédicat à deux places $G(-,-)$.

*Remarque.* Pour alléger l'écriture, on peut convenir d'abréger les lignes génériques lorsqu'elles sont répétées. Par exemple, on peut écrire $\Box$ pour $\forall x(Fx \supset Gxx)$, et $\triangle$ pour $\exists x(Fx \wedge Gxx)$.

### 6.2.3. *Intérêt de la méthode de la preuve par séquents*

A l'opposé des démonstrations faites par la méthode de déduction naturelle, la démonstration par la méthode des séquents conduite conformément au plan de travail est *purement mécanique*. C'est en recherchant méthodiquement à falsifier une formule que l'on découvre les instances particulières de l'axiome $p \Longrightarrow p$ dont on a besoin (dans l'exemple ci-dessus : $Fa \Longrightarrow Fa$ et $Gaa \Longrightarrow Gaa$), ainsi que les formules additionnelles (ou annexes) définies en 4.1.2.2 (respectivement $\forall x(Fx \supset Gxa)$, $Gaa$ et $Fa$) qu'on doit ajouter à ces axiomes, adjonction que l'on justifie en invoquant le principe de monotonie (cf. section 7.7.5).

Aucun effort inventif n'est nécessaire pour trouver les substitutions des variables quantifiées qui révèleront la contradiction, mais on ne peut dire *a priori* quelle sera le cardinal du lexique d'instantiation nécessaire pour prouver la validité d'une formule donnée.

En effet, si on le pouvait, le calcul des prédicats du premier ordre serait *décidable*. On pourrait découvrir mécaniquement en un nombre fini de pas si une formule $\Phi$ de ce calcul est valide, simplement consistante ou inconsistante. Il suffirait d'appliquer la méthode des séquents à cette formule et à sa négation en instanciant avec tous les paramètres du lexique dont la borne supérieure est, dans l'hypothèse envisagée, non seulement finie, mais connue. Or Church a démontré que le calcul des prédicats n'est que semi-décidable.

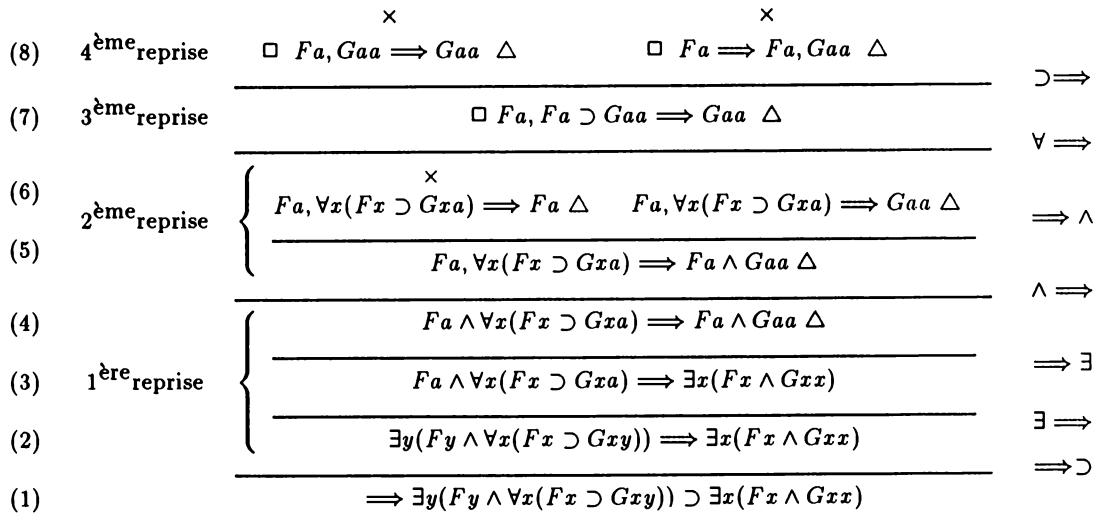



**Figure 6.17** *Preuve par séquents*

### 6.2.4. *Complexité des preuves par séquents*

Reprenons l'exemple 6.10 de la section 6.1.10. On veut prouver par la méthode des séquents :

(6.14) $(\forall x \forall y \forall z((Pxy \wedge Pyz) \supset Gxz) \wedge \forall x \exists y Pyx) \supset \forall x \exists y Gyx.$

Les quatre premières lignes de la démonstration sont les suivantes :

$$\dfrac{\dfrac{\dfrac{(4)\quad \forall x \forall y \forall z((Pxy \wedge Pyz) \supset Gxz), \forall x \exists y Pyx \Longrightarrow \exists y Gya}{(3)\quad \forall x \forall y \forall z((Pxy \wedge Pyz) \supset Gxz) \wedge \forall x \exists y Pyx \Longrightarrow \exists y Gya}\,\wedge\Longrightarrow}{(2)\quad \forall x \forall y \forall z((Pxy \wedge Pyz) \supset Gxz) \wedge \forall x \exists y Pyx \Longrightarrow \forall x \exists y Gyx}\,\Longrightarrow\forall}{(1)\quad \Longrightarrow \forall x \forall y \forall z((Pxy \wedge Pyz) \supset Gxz) \wedge \forall x \exists y Pyx \supset \forall x \exists y Gyx}\,\Longrightarrow\supset$$

On voit tout de suite que de la formule $\forall x \exists y Pxy \Longrightarrow$, *prise isolément*, part une branche qui peut croître indéfiniment, comme dans la figure 6.18.

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{(n+4)\quad \forall x \exists y Pyx, Pba, Pcb \Longrightarrow}}{(n+3)\quad \forall x \exists y Pyx, \exists y Pya, Pba, \exists y Pyb \Longrightarrow}\,\exists\Longrightarrow}{(n+2)\quad \forall x \exists y Pyx, \exists y Pya, Pba \Longrightarrow}\,\forall\Longrightarrow}{(n+1)\quad \forall x \exists y Pyx, \exists y Pya \Longrightarrow}\,\exists\Longrightarrow$$
$$\dfrac{}{(n)\quad \forall x \exists y Pyx \Longrightarrow}\,\forall\Longrightarrow$$

**Figure 6.18** *Branche infinie*

De même, de la formule $\Longrightarrow \exists y Gya$, *prise isolément*, part une branche qui peut croître indéfiniment si on a un lexique $\{a, b, c, \ldots\}$, ce qui est le cas grâce à la formule précédente ($\forall x \exists y Pyx \Longrightarrow$), comme le montre la figure 6.19.

Grâce à l'obligation de puiser dans le lexique $\{a, b, c\}$ dont on dispose en vertu de $\forall x \exists y Pyx \Longrightarrow$, on ne peut instancier que de 27 manières différentes la formule

(6.15) $\forall x \forall y \forall z((Pxy \wedge Pyz) \supset Gxz),$

car 3 objets se prêtent à 27 *permutations avec répétitions*. Au cours de ces instantiations, on tombera inévitablement sur

(6.16) $(Pba \wedge Pcb) \supset Gca.$

$$
\begin{array}{ll}
(m+3) & \Longrightarrow \exists y Gya, Gaa, Gba, Gca \\
\hline
(m+2) & \Longrightarrow \exists y Gya, Gaa, Gba \qquad \Longrightarrow \exists \\
\hline
(m+1) & \Longrightarrow \exists y Gya, Gaa \qquad \Longrightarrow \exists \\
\hline
(m) & \Longrightarrow \exists y Gya
\end{array}
$$

**Figure 6.19** *Branche infinie*

Or on voit tout de suite que le séquent ci-dessous, qui *regroupe* des formules générées à partir d'un lexique de trois paramètres ($a$, $b$, $c$),

(6.17) $(Pba \wedge Pcb) \supset Gca, \quad \underbrace{Pba, Pcb}_{\text{fragment de } n+4} \quad \Longrightarrow \quad \underbrace{Gca}_{\text{fragment de } m+3}$

conduit à l'arbre fermé présenté à la figure 6.20.

$$
\begin{array}{ll}
 & \overset{\times}{Pcb, Pba \Longrightarrow Gca, Pba} \qquad \overset{\times}{Pcb, Pba \Longrightarrow Gca, Pcb} \\
 & \overset{\times}{Gca, Pcb, Pba \Longrightarrow Gca} \qquad \hline \\
(k+1) & \qquad\qquad\qquad\qquad Pcb, Pba \Longrightarrow Gca, (Pcb \wedge Pba) \\
\hline
(k) & (Pcb \wedge Pba) \supset Gca, Pba, Pcb \Longrightarrow Gca
\end{array}
$$

**Figure 6.20** *Arbre fermé*

La rédaction intégrale de cette démonstration aurait été fort longue, mais sa construction est purement mécanique. Nous verrons comment l'extension de la méthode des matrices de Wallen au calcul des prédicats permet de raccourcir sensiblement les preuves conduites selon la méthode des séquents.

### 6.2.5. *Séquents infinis*

Soient deux ensembles de formules $\ldots E_2, E_1, E_0$ et $F_0, F_1, F_2 \ldots$ qui sont des infinités dénombrables (l'un et l'autre ou l'un des deux). On peut leur appliquer le plan de travail moyennant les modifications suivantes (Kleene, 1967, 1972, pp. 312–313) : on isole un sous-ensemble fini dans chacun d'eux à l'aide de marquants, comme indiqué ci-dessous.

(6.18) $\ldots E_{d+2}, E_{d+1}, E_d \mid A_1, \ldots, A_m \Longrightarrow B_1, \ldots B_n \mid F_d, F_{d+1}, F_{d+2} \ldots$

Supposons qu'au cours de la reprise $d$ on ait travaillé sur les formules $A_1, \ldots, A_m$ et $B_1, \ldots, B_n$. A la reprise $d+1$, on déplace les marquants d'un cran vers la droite et vers la gauche, comme suit :

(6.19) $\ldots E_{d+2}, E_{d+1} \mid E_d, A_1, \ldots, A_m \Longrightarrow B_1, \ldots, B_n, F_d \mid F_{d+1}, F_{d+2} \ldots$

Si on découvre un axiome, c'est-à-dire un atome à gauche et à droite de la flèche, mais à l'intérieur des marquants, le chemin se ferme.

> Si *toutes* les formules à l'extérieur des marquants sont des atomes, il y aura terminaison sans fermeture. Une infinité de reprises se déroulera d'une manière instantanée, les marquants étant simplement déplacés de rang en rang (Kleene, *Ibid.*, p. 313).

En vertu du théorème de compacité, qui sera démontré au chapitre 8, un ensemble infini de formules est sémantiquement inconsistant si et seulement si un sous-ensemble fini de celui-ci est sémantiquement inconsistant. Pour prouver l'inconsistance des séquents infinis mentionnées plus haut, il suffira donc de pouvoir montrer l'inconsistance d'un sous-séquent fini du séquent infini initial. En déplaçant les marquants, on accroît constamment la taille du sous-séquent. Cette procédure doit finir par engendrer un arbre fini fermé si toutefois la méthode de preuve par séquents est complète, ce qui est le cas (cf. 8).

### 6.2.6. *Séquents contenant des foncteurs*

Dès qu'il y a des foncteurs, le lexique d'instantiation du séquent est infini, même si le séquent lui-même est fini. Par exemple, dans le séquent

(6.20) $\forall x(\text{Pair } x \supset 2 \text{ divise } x), \text{Pair } (4)^2 \Longrightarrow 2 \text{ divise } (4)^2,$

le lexique auto-généré est l'ensemble

(6.21) $\{2, 4, (2)^2, (4)^2, ((2)^2)^2, ((4)^2)^2 \ldots\}.$

On instanciera donc successivement $\forall x$ (Pair $x \supset 2$ divise $x$) à l'aide des constantes figurant dans le lexique infini ci-dessus. Arrivé à l'instance $(4)^2$, l'arbre se ferme. Un moment de réflexion nous fera choisir l'instance $(4)^2$ d'emblée, mais même sans appel à la réflexion, l'exploration systématique du lexique nous aurait conduit *mécaniquement* à l'arbre de preuve fini fermé.

La méthode des séquents peut également s'appliquer à des ensembles infinis de formules dans lesquels figurent des ensembles infinis de variables libres, de constantes et de symboles de fonctions (cf. Kleene 1969, 1972, pp. 313–315).

## 6.3. Matrices de Bibel et de Wallen

La méthode exposée ici est spécialement conçue pour les applications sur ordinateur. La section 6.3 peut être omise en première lecture.

### 6.3.1. *Non-déterminisme de la preuve par séquents*

La procédure de preuve de Kleene, avec son plan de travail et son lexique d'instantiation autogénéré, conduit à une preuve du séquent d'entrée d'une

manière "presque complètement déterminée". Comme l'écrit Robinson,

> [...] La seule place pour un choix réside dans les séquences "passives" des cycles, là où il y a du jeu (*there is leeway*) quant à la question de savoir quelles formules passives il convient de choisir (Robinson, 1979, p. 115).

A chaque reprise, on fait un pas avec *chacune* des formules passives (de type $\alpha$, $\beta$, $\delta$) et un nombre *fini* de pas avec les formules actives ($\gamma$), mais Kleene ne nous dit pas dans quel *ordre* il faut prendre les formules passives figurant parmi les $A_1, \ldots, A_m$, $B_1, \ldots, B_n$ d'un séquent, c'est-à-dire d'une expression de la forme

(6.22) $\quad A_1, \ldots, A_m \Longrightarrow B_1, \ldots, B_n.$

Certes, on peut se borner à suivre *l'ordre linéaire de lecture du séquent*, mais on s'expose alors aux trois critiques formulées par Wallen à la méthode des séquents pour le calcul des propositions : la *redondance*, la *non-pertinence* et la *complexité* (cf. 4.2.1). Il existe une méthode qui a les avantages de la méthode des séquents sans en avoir les inconvénients. Bien qu'elle soit économique et facile à employer, elle est assez difficile à exposer. C'est la méthode de Bibel et de Wallen, qui est non seulement *mécanique* mais, par surcroît, *déterministe*.

### 6.3.2. *Calcul des prédicats et connexions*

Pour étendre la méthode des connexions au calcul des prédicats, il faudra commencer par ajouter aux règles $\alpha$ et $\beta$ (4.2.5) des règles de *démembrement syntaxique* pour les formules de type $\gamma$ et de type $\delta$. Ces règles sont données à la figure 6.21.

| $\gamma$ | $\gamma_0(a)$ |
|---|---|
| $\langle \forall x A, \mathbf{V} \rangle$ | $\langle Aa/x, \mathbf{V} \rangle$ |
| $\langle \exists x A, \mathbf{F} \rangle$ | $\langle Aa/x, \mathbf{F} \rangle$ |

| $\delta$ | $\delta_0(a)$ |
|---|---|
| $\langle \forall x A, \mathbf{F} \rangle$ | $\langle Aa/x, \mathbf{F} \rangle$ |
| $\langle \exists x A, \mathbf{V} \rangle$ | $\langle Aa/x, \mathbf{V} \rangle$ |

**Figure 6.21** *Règles $\gamma$, $\delta$*

### 6.3.3. *Le chemin en calcul des prédicats*

On ajoute les clauses suivantes à la définition récursive des chemins :

4. Si $S$ est un chemin contenant le noeud $\gamma^k$, alors $S, \gamma_0^{kj}$ est un chemin pour $1 \leq j \leq \mu(\gamma_0)$, où $\mu$ désigne le nombre d'instantiations de la formule active $\gamma$. Ce nombre est déterminé ainsi : on prolonge la formule-mère $\gamma$ par autant de formules-filles qu'il faut pour arriver à une *connexion* dans chaque chemin, en respectant les contraintes imposées à $\gamma^{kj}$ (appartenir au lexique et ne pas avoir déjà été utilisé sur la même branche) et à $\delta^{kj}$ (être inédite et être "fille unique", c'est-à-dire ne pas avoir de "soeur").

   Remarquons que l'on n'élimine pas la formule-mère. Cette particularité de la règle pour les formules de type $\gamma$ a déjà été observée dans le *calcul des séquents* (les quantificateurs universels ou quasi-universels sont répétés à côté de leurs instances) ainsi que dans la *méthode des tableaux* (les quantificateurs universels ou quasi-universels ne sont pas cochés).
5. Si $S, \delta^k$ est un chemin, alors $S \backslash \{\delta^k\}, \delta_0^k$ est également un chemin; en d'autres termes, on prolonge le chemin en éliminant la formule-mère $\delta^k$ et en la remplaçant par la formule-fille $\delta_0^k$.

### 6.3.4. *La notion de muliplicité*

Dans le cas des formules $\gamma$, il y a, en principe, une infinité de descendants $\gamma_0$ puisqu'il y a une infinité de paramètres (de variables d'instantiation $a_i$). Cependant, le théorème de compacité nous assure que si une formule ou un ensemble de formules sont sémantiquement inconsistants, cela peut se prouver à l'aide d'une preuve finie, faisant donc intervenir un nombre fini d'instances des formules $\gamma$ : $\gamma_0^1, \ldots, \gamma_0^n$. Nous pouvons donc prolonger — dans l'arbre syntaxique — chaque formule $\gamma$ par $n$ descendants. Comme nous l'avons vu (section 6.2.3), on ne peut pas décider *a priori* quel est le cardinal du lexique dont on aura besoin pour démontrer l'inconsistance d'une formule ou d'un ensemble de formules. La bonne technique c'est d'essayer d'abord une *multiplicité* d'instances $\gamma_0$ de cardinal 1, puis — si on n'a pas réussi à trouver une preuve — d'accroître la multiplicité d'une unité et ainsi de suite.

La multiplicité[5] d'instances d'une forme $X$ de type $\gamma$ sera indiquée par la fonction $\mu$ qui prend ses arguments dans les formules $X$ de type $\gamma$ et ses valeurs dans $\mathbf{N}$. Par exemple, dans la preuve par séquents examinée en 6.2.2, on a

(6.23) $\quad \mu[\forall x(Fx \supset Gxa)] = 1.$

Pour chaque formule $X$ de type $\gamma$, le chemin qui la traverse doit comporter $\mu$ instances $\gamma$ de $X$. Concrètement, $\mu$ est "la mesure de l'effort de création d'instances" que l'on demande aux formules créatives de type $\gamma$.

[5] En 6.3.11, on verra un cas où la multiplicité $\mu$ est supérieure à 1.

### 6.3.5. *Comment définir l' inédit ?*

Dans le calcul des propositions, on descend dans l'*arbre syntaxique* jusqu'aux atomes. On néglige totalement la structure interne des *arbres de preuve*. On se préoccupe seulement des *atomes* terminaux qui apparaissent aux extrémités des branches de l'arbre syntaxique. Le *type* de l'atome est important car la position de l'atome dans l'arbre des chemins en dépend. La *polarité* aussi est importante. En effet, pour obtenir une "connexion", on doit arriver à deux atomes identiques et de polarité opposée sur un même chemin. *Dans le calcul des prédicats, on ne peut plus ignorer la structure interne de la dérivation en calcul des séquents.* En effet, la dérivation par séquents est soumise à la contrainte suivante : les paramètres qui interviennent dans l'instantiation des formules de type $\delta$ doivent être *inédits*.

Si l'on fait l'économie de la construction d'un *arbre de preuve* et que l'on se borne à utiliser un *arbre syntaxique* pour accéder aux atomes, ce qui était l'innovation majeure de Wallen, comment pourra-t-on reformuler cette contrainte ? La tâche est d'autant plus difficile que la notion de "paramètre inédit" est relative à l'arbre de preuve : tel paramètre inédit à la ligne $n$ d'une dérivation peut cesser de l'être à la ligne $n+1$. Nous allons voir, cependant, que la difficulté n'est pas insurmontable.

### 6.3.6. *Alignement des universelles sur les existentielles*

Pour générer des atomes complémentaires sur un chemin, tout en respectant la contrainte qui pèse sur les instances de propositions de type $\delta$ (instances qui doivent être inédites), le *plan de travail* en vigueur dans le calcul des séquents prévoyait ce qui suit : on donnait la priorité aux formules passives (notamment aux formules $\delta$) et, pour instantier les formules $\gamma$, on *réutilisait* les paramètres introduits lors de l'instantiation des formules $\delta$. Par exemple, pour établir l'inconsistance de $\{\forall x \neg Fx, \exists x Fx\}$, on commençait par instantier $\exists x Fx$ en $Fa$ et ensuite on instantiait $\forall x \neg Fx$ en $\neg Fa$, en réutilisant le paramètre $a$. En d'autre termes, on *alignait* les instances $\gamma_0$ de $\gamma$ sur les instances $\delta_0$ de $\delta$.

Wallen utilise les avantages respectifs de la méthode des connexions et de la méthode de résolution : la décomposition unique de la formule en sous-formules et l'existence d'un algorithme d'unification des variables d'instantiation. Pour cela, il utilise comme variables d'instantiation[6] les pointeurs de position dans l'arbre syntaxique $\{a_0, a_1, \ldots\}$ munies d'un exposant $(\mu)$ rappelant que la sous-formule a pour ancêtre une formule $\gamma$ instanciée un nombre $\mu$ de fois. Cela conduit à un *arbre de formule indexé* pour chaque choix de multiplicité.

[6] Nous parlons de variables d'instantiation plutôt que de paramètres parce que nous allons devoir opérer des unifications entre ces dernières. Or l'unification est une forme de substi-

Ensuite, il parcourt chaque chemin pour y découvrir des connexions. Les atomes passés en revue peuvent comporter des variables d'instantiation provenant de l'application des règles $\gamma$ et $\delta$. Ces variables constituent les deux ensembles $\Gamma_0(\mu)$ et $\Delta_0(\mu)$. Pour que deux atomes soient complémentaires (forment une connexion), il faut trouver une substitution $\sigma$ des variables d'instantiation d'universelles $\Gamma_0$ vers les variables d'instantiation d'universelles ou d'existentielles $\Gamma_0 \cup \Delta_0$ qui unifie les instances et qui respecte la contrainte requérant des variables d'*instantiation existentielle* qui soient *inédites* (comme dans la méthode des tableaux et dans celle des séquents). Ces deux opérations : *trouver une substitution qui unifie* des atomes complémentaires dans chaque chemin et *respecter* la contrainte d'ordre sont ici bien séparées. Une fois la substitution $\sigma$ découverte par algorithme, on vérifiera que la substitution respecte la contrainte de l'ordre d'instantiation.

### 6.3.7. *L'ordre de réduction des formules*

La relation $\lhd$ est assujettie à la contrainte unique suivante : elle doit être *irréflexive*. Cette relation est définie à l'aide de l'union de deux autres relations : la relation *a pour sous-formule* ($\ll$) et la relation *est instancié avant* ($\sqsubset$), qu'il convient de définir.

**Définition 6.1 : la relation $\ll$.** L'état des choses $A \ll B$ est réalisé si et seulement si $B$ est sous-formule de $A$ dans l'arbre syntaxique.

**Définition 6.2 : la relation $\sqsubset$.** Soit $\Gamma_0$ l'ensemble des filles d'universelles et $\Delta_0$ l'ensemble des filles d'existentielles. La fonction $\sigma : \Gamma_0(\mu) \mapsto \Gamma_\mu \cup \Delta(\mu)$ induit la relation d'ordre $\sqsubset$ définie sur (le produit cartésien) $\Delta_0(\mu) \times \Gamma_0(\mu)$ et la relation d'équivalence $\approx$ définie sur (le produit cartésien) $\Gamma_0(\mu) \times \Gamma_0(\mu)$.

Les trois conditions suivantes sont satisfaites (Wallen, 1987, p. 81) :

1. Si $\sigma(u) = v$ et $v \in \Gamma_0(\mu)$, alors $u \approx v$.
2. Si $\sigma(u) = v$ et $v \in \Delta_0(\mu)$, alors $v \sqsubset u$.
3. Si $v \sqsubset u$ et $u \approx u'$, alors $v \sqsubset u'$.

Par exemple, si $A$ est la position du pointeur d'une formule de type $\gamma_0$, et $B$ le pointeur d'une formule de type $\delta_0$, on aura

$$\sigma : A \mapsto B$$

et

$$B \sqsubset A,$$

---

tution et on ne peut effectuer de substitution sur un paramètre, celui-ci ayant le statut de constante.

qui peut se lire "$B$ est instancié *avant* $A$". Comme nous l'avons vu plus haut, on utilise le pointeur (indicateur de position) de la formule en guise de variable d'instantiation.

**Définition 6.3 : la relation $\lhd$.** La relation $\lhd$ se définit ainsi, sachant que le signe $+$ placé en exposant désigne la fermeture transitive :

$$\lhd =_{def} (\ll \cup \sqsubset)^{+}.$$

### 6.3.8. *Les étapes d'une preuve par la méthode des connexions*

**Phase 1 :** on construit l'arbre syntaxique de la formule en appliquant le principe de parcours : "en profondeur d'abord avec retour en arrière".

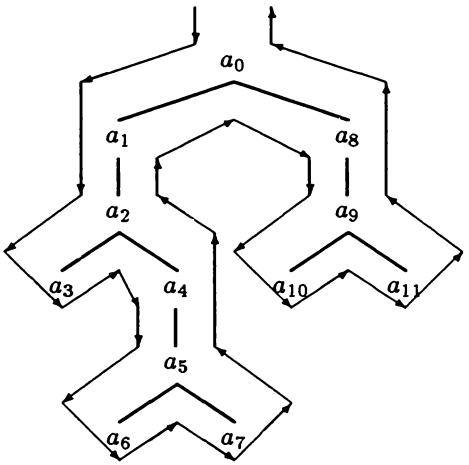



**Figure 6.22** *Le principe "en profondeur d'abord"*

**Phase 2 :** on reprend dans un arbre de formule indexé tous les constituants obtenus au cours de l'analyse syntaxique de la formule. Chaque constituant reçoit dans la colonne $u$ un numéro d'ordre qui est celui du noeud auquel il est associé $(a_0, \ldots, a_j)$. On assigne un type principal et un type secondaire à chaque constituant. Enfin — en s'appuyant sur le type et sur les règles $\alpha, \beta, \gamma, \delta$ — on assigne une polarité à chaque constituant. La polarité de la formule principale est 0 puisqu'on essaie de la falsifier.

**Phase 3 :** à partir du tableau construit dans la phase 2, et en appliquant les clauses de la définition récursive du chemin, on construit les chemins qui conduisent à isoler des collections d'atomes dans lesquelles on va chercher à identifier la présence de *connexions*, c'est-à-dire d'atomes complémentaires.

**Phase 4 :** chercher dans chaque chemin des atomes complémentaires moyennant substitution ("atomes $\sigma$-complémentaires"). Les paramètres de chaque constituant sont chaque fois ***différents***. On a, en effet, pris comme paramètre d'instantiation le nom du noeud du constituant considéré; on doit ensuite ***uniformiser*** ces derniers en ***alignant*** les paramètres d'un atome sur celui d'un autre atome tout en ***respectant*** les contraintes. Notons que les noms de noeuds de propositions de type $\gamma_0$ sont surmontés d'un trait horizontal qui permet de les distinguer des $\delta_0$.

**Phase 5 :** on construit le graphe de la relation $\lhd$ en représentant par des flèches pleines la relation $\ll$ et par des flèches en pointillés la relation $\sqsubset$. Pour que la preuve de validité de la formule soit recevable, la relation $\lhd$ doit être irréflexive. L'introduction de flèches en pointillés transforme l'*arbre* de preuve en *graphe*. Exiger que la relation $\lhd$ soit irréflexive, c'est exiger que le graphe en question soit acyclique, c'est-à-dire qu'il ne contienne aucune boucle.

Cette méthode que Wallen a étendue aux logiques modales et à la logique intuitionniste repose sur la définition suivante de la validité due à Bibel :

> Une formule $F$ est valide si et seulement si pour quelque multiplicité $\mu$, il existe un ensemble $U$ de connexions couvrant tous les chemins (***spanning set***) dans $F^\mu$ et une substitution $\sigma$ telle que $u\sigma$ est une paire de littéraux complémentaires pour n'importe quel $u \in U$ et la relation $\lhd$ déterminée par $\sigma$ et $F^\mu$ ne comporte aucun cycle (Bibel, 1982, p. 185).

### 6.3.9. *Exemple d'une preuve par la méthode des connexions*

Nous allons montrer qu'on ne peut trouver de contre-exemple à la proposition

(6.24) S'il y a un philosophe que tous les philosophes contredisent, alors il y a un philosophe qui se contredit lui-même..

Pour cet exercice, la multiplicité $\mu$ pour les formules créatives ($\gamma$) est 1.

**Phase 1 :** *Construction de l'arbre*

L'arbre syntaxique de la proposition est présenté à la figure 6.23.

**Phase 2 :** *Construction de l'arbre de formule indexé*

nous allons à présent construire un tableau formé de cinq colonnes. Dans la première ($u$), nous énumérons les noms des nœuds. Ces noms sont

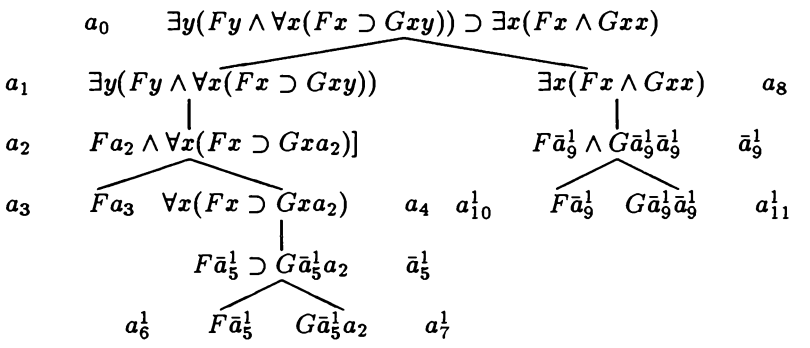



**Figure 6.23** *Arbre syntaxique de la formule 6.24*

des lettres $a_i$ où la succession des $i$ retrace le parcours du graphe. Dans la deuxième (pol($u$)) figure la polarité (**V**[1] ou **F**[0]) de chacune des formules inscrites dans la troisième colonne (lab($u$)). Ces formules sont tous les constituants ou sous-formules de la formule initiale, y compris celle-ci. Dans la quatrième colonne (Ptype($u$)), on écrit le *type principal* de chaque formule en tant que mère. Les formules atomiques n'ont pas de type principal puisqu'elles n'ont pas de descendant. Dans la cinquième colonne, on écrit le *type secondaire* de chaque formule, c'est-à-dire le type auquel appartient la formule en tant que fille. La formule initiale n'a dès lors pas de type secondaire (cf. figure 6.24).

| $u$ | pol($u$) | lab($u$) | Ptype($u$) | Stype($u$) |
|---|---|---|---|---|
| $a_0$ | 0 | $\exists y(Fy \wedge \forall x(Fx \supset Gxy)) \supset \exists x(Fx \wedge Gxx)$ | – | $\alpha$ |
| $a_1$ | 1 | $\exists y(Fy \wedge \forall x(Fx \supset Gxy))$ | $\delta$ | $\alpha_1$ |
| $a_2$ | 1 | $Fa_2 \wedge \forall x(Fx \supset Gxa_2)$ | $\alpha$ | $\delta_0$ |
| $a_3$ | 1 | $Fa_2$ | – | $\alpha_1$ |
| $a_4$ | 1 | $\forall x(Fx \supset Gxa_2)$ | $\gamma$ | $\alpha_2$ |
| $\bar{a}_5^1$ | 1 | $F\bar{a}_5^1 \supset G\bar{a}_5^1 a_2$ | $\beta$ | $\gamma_0$ |
| $a_6^1$ | 0 | $F\bar{a}_5^1$ | – | $\beta_1$ |
| $a_7^1$ | 1 | $G\bar{a}_5^1 a_2$ | – | $\beta_2$ |
| $a_8$ | 0 | $\exists x(Fx \wedge Gxx)$ | $\gamma$ | $\alpha_2$ |
| $\bar{a}_9^1$ | 0 | $F\bar{a}_9^1 \wedge G\bar{a}_9^1\bar{a}_9^1$ | $\beta$ | $\gamma_0$ |
| $a_{10}^1$ | 0 | $F\bar{a}_9^1$ | – | $\beta_1$ |
| $a_{11}^1$ | 0 | $G\bar{a}_9^1\bar{a}_9^1$ | – | $\beta_2$ |

**Figure 6.24** *Arbre de formule indexé*

On instancie les quantificateurs à l'aide des lettres numérotées figurant dans la colonne $u$. Ces lettres numérotées ont le statut logique de variables. Celles d'entre elles qui servent de variables d'instantiation à des quantificateurs *universels* ou *quasi-universels* sont surmontées d'un trait horizontal et d'un indice $t_j$, pour $1 \leq j \leq \mu(\gamma_0)$.

**Phase 3 :** *Construction des chemins*

| | | | | |
|---|---|---|---|---|
| 1 | | $\{a_0\}$ | | |
| 2 | | $\{a_1, a_8\}$ | | |
| 3 | | $\{a_2, a_8, \bar{a}_9^1\}$ | | |
| 4 | | $\{a_3, a_4, a_8, \bar{a}_9^1\}$ | | |
| 5 | $\{a_3, a_4, a_8, a_{10}^1\}$ | | $\{a_3, a_4, a_8, a_{11}^1\}$ | |
| 6 | $\{a_3, a_4, \bar{a}_5^1, a_8, a_{10}\}$ | | $\{a_3, a_4, \bar{a}_5^1, a_8, a_{11}^1\}$ | |
| 7 | $\{a_3, a_4, a_6^1, a_8, a_{10}^1\}$ | $\{a_3, a_4, a_7^1, a_8, a_{10}^1\}$ | $\{a_3, a_4, a_6^1, a_8, a_{11}^1\}$ | $\{a_3, a_4, a_7^1, a_8, a_{11}^1\}$ |

**Figure 6.25** *Chemins*

Remarquons que les formules-mères de type $\gamma$, une fois apparues dans un chemin, y restent. Dans l'exemple présenté à la figure 6.25, c'est le cas des formules désignées par $a_4$ ($\mathbf{V}\forall x A$) et $a_8$ ($\mathbf{F}\exists x a$).

**Phase 4 :** *Connexions*

Dans chaque chemin on trouve deux littéraux que l'on peut rendre complémentaires par substitution $\sigma$, respectivement $\{a_3, a_6^1\}$, $\{a_3, a_{10}^1\}$, $\{a_3, a_6^1\}$, $\{a_7^1, a_{11}^1\}$, c'est-à-dire :
$\{\langle \mathbf{V}, Fa_2\rangle, \langle \mathbf{F}, F\bar{a}_5^1\rangle\}$, $\{\langle \mathbf{V}, Fa_2\rangle, \langle \mathbf{F}, F\bar{a}_9\rangle\}$,
$\{\langle \mathbf{V}, Fa_2\rangle, \langle \mathbf{F}, F\bar{a}_5^1\rangle\}$, $\{\langle \mathbf{V}, G\bar{a}_5^1 a_2\rangle, \langle \mathbf{F}, G\bar{a}_9^1\bar{a}_9^1\rangle\}$.

$$\sigma : \overline{a}_5^1 \mapsto a_2 \qquad \sigma : \overline{a}_9^1 \mapsto a_2$$

**Phase 5 :** *Construction du graphe*

A la relation d'ordre (*a pour sous-formule*) on combine les relations d'ordre (*est instanciée avant*) obtenues à partir de $\sigma$ : $a_2 \sqsubset \overline{a}_5^1, a_2 \sqsubset \overline{a}_9^1$.

Le graphe est présenté à la figure 6.26. Rappelons que dans le graphe (où $a_i$ est au-dessus de $a_j$) $a_i \rightarrow a_j$ signifie que $a_i$ a pour sous-formule $a_j$, et que $a_i \cdots > a_j$ signifie que $a_i$ *précède* $a_j$. Le graphe reste acyclique : il ne contient pas de boucle.

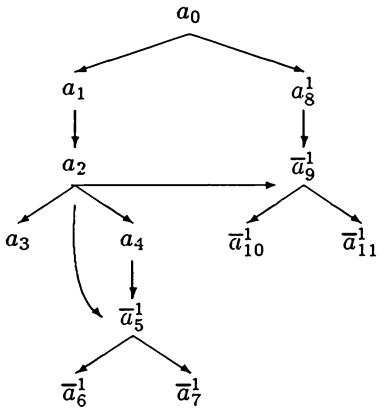



**Figure 6.26** *Graphe acyclique*

### 6.3.10. *Les mérites de la méthode des connexions*

Tous les avantages que la méthode de Wallen possédait dans le calcul des propositions se retrouvent ici (absence de redondance, mise hors jeu des traits

non pertinents, caractère déterministe et unique de l'arbre de preuve). A ceux-ci s'en ajoute un nouveau : l'unification des paramètres peut être faite en appliquant un algorithme tel que l'*algorithme d'unification* de J. Robinson, algorithme qui sera décrit en 6.4.10. Comme l'écrit Wallen :

> Dans une recherche fondée sur la méthode des matrices, nous n'avons pas besoin de choisir les paramètres particuliers, ni de choisir un ordre particulier d'application des règles des séquents. Nous nous préoccupons seulement d'inférer l'indispensable cohérence que tout choix de paramètres doit posséder, et de nous assurer qu'il existe *au moins* un ordre de réduction [de la formule à ses atomes] qui respecte les contraintes pesant sur les règles gouvernant les quantificateurs (Wallen, 1987, p. 81).

Remarquons que l'arbre de preuve — qui est aussi un arbre syntaxique de la formule — est *unique* pour une multiplicité donnée. Tout travail de création se borne à choisir la multiplicité de chaque universelle.

### 6.3.11. *Application*

Considérons à nouveau la formule traitée en 6.1.10

(6.25) $(\forall x \exists y Pyx \wedge \forall x \forall y \forall z((Pxy \wedge Pyz) \supset Gxz)) \supset \forall x \exists y Gyx.$

**Phase 1 :** construction de l'arbre syntaxique.

La figure 6.27 présente l'arbre syntaxique associé à la formule 6.25.

**Phase 2 :** choix des multiplicités et construction de l'arbre de formule indexé. La figure 6.28 présente l'arbre syntaxique sous une forme schématique (les pointeurs $a_i$ remplacent les formules et sous-formules), mais en indiquant la multiplicité.

Lorsqu'il y a plusieurs quantificateurs universels emboîtés — ce qui n'était pas le cas dans l'exemple 6.24 — il est utile de faire usage d'exposants. $\overline{a}_i^{1\,1}$ signifie que la formule désignée par $\overline{a}_i$ est de type $\gamma_0$ (puisqu'elle est surmontée d'un trait horizontal) et que la formule-mère $\gamma$ était elle-même de type secondaire $\gamma_0$. Plus simplement, le *nombre* d'exposants — ici il y en a deux — indique le nombre de quantificateurs universels que l'on a éliminés pour arriver à $\overline{a}_i$. Si on avait écrit $\overline{a}_i^{1\,2}$, le nombre d'exposants aurait été le même — il y aurait eu deux exposants — mais le chiffre '2' du deuxième exposant aurait indiqué que la formule considérée était la *deuxième instance* d'une même formule-mère $\overline{a}_i^1$, ce qui suppose que $\mu(\overline{a}_i^{1\,2})$ est *au moins* égal à 2.

[7] Pour alléger l'écriture, le symbole $\wedge$ entre $Pxy$ et $Pyz$ dans la figure 6.27 est sous-entendu à partir de la deuxième ligne.

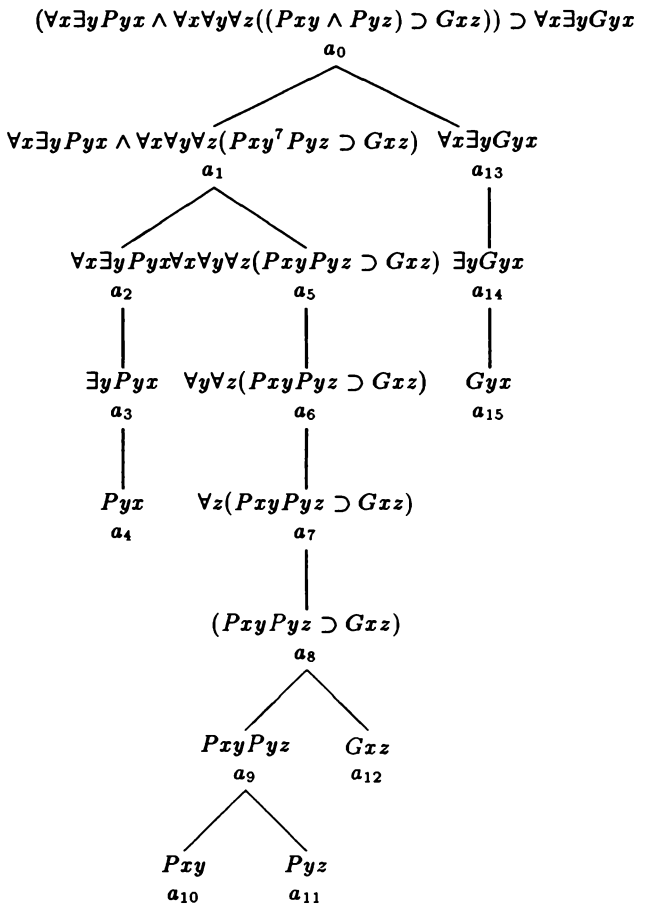
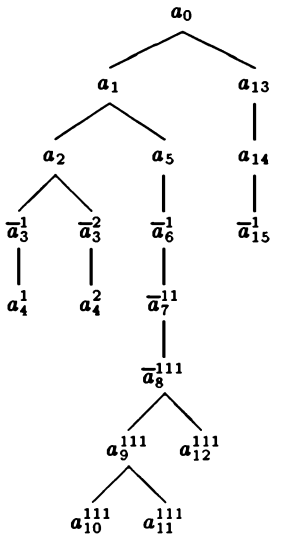



**Figure 6.27** *Arbre syntaxique de 6.25*

$\mu(\overline{a}_3) = 2,$
$\mu(\overline{a}_6) = \mu(\overline{a}_7) = \mu(\overline{a}_8) = \mu(\overline{a}_{15}) = 1,$
$\Gamma_0 = \{\overline{a}_3^1, \overline{a}_3^2, \overline{a}_6^1, \overline{a}_7^{11}, \overline{a}_8^{111}, \overline{a}_{15}^1\}$
($\Gamma_0$ est l'ensemble des filles d'universelles),
$\Delta_0 = \{a_4^1, a_4^2, a_{14}\}$
($\Delta_0$ est l'ensemble des filles d'existentielles)

On verra plus loin que si on avait adopté une multiplicité moindre pour $\overline{a}_3$, à savoir $\mu(\overline{a}_3) = 1$, on aurait obtenu un graphe cyclique. On aurait donc échoué à établir la validité de la formule valide 6.25.

La figure 6.29 présente l'arbre de formule indexé.




La bifurcation partant de $a_2$ correspond à la multiplicité $\mu[(\forall x)(\exists y)Pyx] = 2$. si on avait eu $\mu = 3$, la bifurcation aurait été remplacée par un trident.

**Figure 6.28** *Multiplicités*

**Phase 3 :** construction de chemins à partir de l'arbre de formule indexé , par application des clauses de la définition récursive de "chemin".

**Phase 4 :** chercher dans chaque chemin des atomes $\sigma$-complémentaires (c'est-à-dire complémentaires moyennant substitution). On isole les atomes dans chaque feuille terminale de l'arbre (ligne 12, figure 6.30).

- Chemin $1 = \{a_4^1, a_4^2, a_{10}^{111}, \overline{a}_{15}^1\}$, c'est-à-dire
  $\{\mathbf{V}Pa_4^1\overline{a}_3^1, \underline{\mathbf{V}Pa_4^2\overline{a}_3^2}, \underline{\mathbf{F}P\overline{a}_6^1\overline{a}_7^{11}}, \mathbf{F}G\overline{a}_{15}^1a_{14}\}$
- Chemin $2 = \{a_4^1, a_4^2, a_{11}^{111}, \overline{a}_{15}^1\}$, c'est-à-dire
  $\{\underline{\mathbf{V}Pa_4^1\overline{a}_3^1}, \mathbf{V}Pa_4^2\overline{a}_3^2, \underline{\mathbf{F}P\overline{a}_7^{11}\overline{a}_8^{111}}, \mathbf{F}G\overline{a}_{15}^1a_{14}\}$
- Chemin $3 = \{a_4^1, a_4^2, a_{12}^{111}, \overline{a}_{15}^1\}$, c'est-à-dire
  $\{\mathbf{V}Pa_4^1\overline{a}_3^1, \mathbf{V}Pa_4^2\overline{a}_3^2, \underline{\mathbf{V}G\overline{a}_6^1\overline{a}_8^{111}}, \underline{\mathbf{F}G\overline{a}_{15}^1a_{14}}\}$

Procéder à l'unification des atomes soulignés.

Résultat de l'unification :

- Chemin 1 : $Pa_4^2\overline{a}_3^2$ et $P\overline{a}_6^1\overline{a}_7^{11}$ donnent $Pa_4^2\overline{a}_3^2$
- Chemin 2 : $Pa_4^1\overline{a}_3^1$ et $P\overline{a}_7^{11}\overline{a}_8^{111}$ donnent $Pa_4^1\overline{a}_3^1$
- Chemin 3 : $G\overline{a}_6^1\overline{a}_8^{111}$ et $G\overline{a}_{15}^1a_{14}$ donnent $G\overline{a}_6^1a_{14}$

| $\mu$ | pol($\mu$) | lab($\mu$) | Ptype($\mu$) | Stype($\mu$) |
|---|---|---|---|---|
| $a_0$ | 0 | $(\forall x \exists y Pyx \wedge \forall x \forall y \forall z((Pxy \wedge Pyz) \supset Gxz)) \supset \forall x \exists y Gyx$ | $\alpha$ | – |
| $a_1$ | 1 | $\forall x \exists y Pyx \wedge \forall x \forall y \forall z((Pxy \wedge Pyz) \supset Gxz)$ | $\alpha$ | $\alpha_1$ |
| $a_2$ | 1 | $\forall x \exists y Pyx$ | $\gamma$ | $\alpha_1$ |
| $\overline{a}_3^1$ | 1 | $\exists y Py\overline{a}_3^1$ | $\delta$ | $\gamma_0$ |
| $\overline{a}_3^2$ | 1 | $\exists y Py\overline{a}_3^2$ | $\delta$ | $\gamma_0$ |
| $a_4^1$ | 1 | $Pa_4^1\overline{a}_3^1$ | – | $\delta_0$ |
| $a_4^2$ | 1 | $Pa_4^2\overline{a}_3^2$ | – | $\delta_0$ |
| $a_5$ | 1 | $\forall x \forall y \forall z((Pxy \wedge Pyz) \supset Gxz)$ | $\gamma$ | $\alpha_2$ |
| $\overline{a}_6^1$ | 1 | $\forall y \forall z((P\overline{a}_6^1 y \wedge Pyz) \supset G\overline{a}_6^1 z)$ | $\gamma$ | $\gamma_0$ |
| $\overline{a}_7^{11}$ | 1 | $\forall z((P\overline{a}_6^1\overline{a}_7^{11} \wedge P\overline{a}_7^{11} z) \supset G\overline{a}_6^1 z)$ | $\gamma$ | $\gamma_0$ |
| $\overline{a}_8^{111}$ | 1 | $(P\overline{a}_6^1\overline{a}_7^{11} \wedge P\overline{a}_7^{11}\overline{a}_8^{111}) \supset G\overline{a}_6^1\overline{a}_8^{111}$ | $\beta$ | $\gamma_0$ |
| $a_9^{111}$ | 0 | $P\overline{a}_6^1\overline{a}_7^{11} \wedge P\overline{a}_7^{11}\overline{a}_8^{111}$ | $\beta$ | $\beta_1$ |
| $a_{10}^{111}$ | 0 | $P\overline{a}_6^1\overline{a}_7^{11}$ | – | $\beta_1$ |
| $a_{11}^{111}$ | 0 | $P\overline{a}_7^{11}\overline{a}_8^{111}$ | – | $\beta_2$ |
| $a_{12}^{111}$ | 1 | $G\overline{a}_6^1\overline{a}_8^{111}$ | – | $\beta_2$ |
| $a_{13}$ | 0 | $\forall x \exists y Gyx$ | $\delta$ | $\alpha_2$ |
| $a_{14}$ | 0 | $\exists y Gya_{14}$ | $\gamma$ | $\delta_0$ |
| $\overline{a}_{15}^1$ | 0 | $G\overline{a}_{15}^1 a_{14}$ | – | $\gamma_0$ |

**Figure 6.29** *Arbre de formule indexé*

L'unification a été faite grâce à la substitution $\sigma = \langle \overline{a}_6^1, a_4^2 \rangle, \langle \overline{a}_7^{11}, \overline{a}_3^2 \rangle, \langle \overline{a}_7^{11}, a_4^1 \rangle, \langle \overline{a}_8^{111}, \overline{a}_3^1 \rangle, \langle \overline{a}_{15}^1, \overline{a}_6^1 \rangle, \langle \overline{a}_8^{111}, a_{14} \rangle$.

**Phase 5 :** on construit le graphe de la relation d'ordre $\lhd = (\ll \cup \sqsubset)^+$ où $\ll$ signifie la relation *a pour sous-formule*, représentée par le graphe de la figure 6.28 lu de haut en bas, et où l'autre relation d'ordre $\sqsubset$ représente la relation *est instancié avant*. On construit la relation $\sqsubset$ à partir de $\sigma$ en respectant les règles de 6.2, à savoir :

1. si $\sigma(\overline{a}_i) = \overline{a}_j$, alors $\overline{a}_i \approx \overline{a_j}$ (pas d'ordre entre les formules $\gamma$)
2. si $\sigma(\overline{a}_i) = a_k$, alors $a_k \sqsubset \overline{a_i}$ (instancier $\delta$ *avant* $\gamma$)
3. si $a_k \sqsubset \overline{a_i}$ et $\overline{a}_i \approx \overline{a_j}$, alors $a_k \sqsubset \overline{a_j}$

Cela donne :

- en vertu de (1), on a

$$\overline{a}_7^{11} \approx \overline{a}_3^2$$
$$\overline{a}_8^{111} \approx \overline{a}_3^1$$
$$\overline{a}_{15}^1 \approx \overline{a}_6^1$$

| | | | |
|---|---|---|---|
| 1 | | $\{a_0\}$ | |
| 2 | | $\{a_1, a_{13}\}$ | |
| 3 | | $\{a_2, a_5, a_{13}\}$ | |
| 4 | | $\{a_2, \overline{a}_3^1, \overline{a}_3^2, a_5, a_{13}\}$ | |
| 5 | | $\{a_2, a_4^1, a_4^2, a_5, a_{13}\}$ | |
| 6 | | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, a_{13}\}$ | |
| 7 | | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, \overline{a}_7^{11}, a_{13}\}$ | |
| 8 | | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, \overline{a}_7^{11}, \overline{a}_8^{111}, a_{13}\}$ | |
| 9 | | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, \overline{a}_7^{11}, a_9^{111}, a_{13}\}$ | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, \overline{a}_7^{11}, a_{12}^{111}, a_{13}\}$ |
| 10 | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, \overline{a}_7^{11}, a_{10}^{111}, a_{13}\}$ | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, \overline{a}_7^{11}, a_{11}^{111}, a_{13}\}$ | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, \overline{a}_7^{11}, a_{12}^{111}, a_{13}\}$ |
| 11 | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, \overline{a}_7^{11}, a_{10}^{111}, a_{14}\}$ | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, \overline{a}_7^{11}, a_{11}^{111}, a_{14}\}$ | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, \overline{a}_7^{11}, a_{12}^{111}, a_{14}\}$ |
| 12 | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, \overline{a}_7^{11}, a_{10}^{111}, a_{14}, \overline{a}_{15}^1\}$ | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, \overline{a}_7^{11}, a_{11}^{111}, a_{14}, \overline{a}_{15}^1\}$ | $\{a_2, a_4^1, a_4^2, a_5, \overline{a}_6^1, \overline{a}_7^{11}, a_{12}^{111}, a_{14}, \overline{a}_{15}^1\}$ |

**Figure 6.30** *Chemins pour la formule 6.25*

- en vertu de (2), on a

$$a_4^2 \sqsubset \overline{a}_6^1$$
$$a_4^1 \sqsubset \overline{a}_7^{11}$$
$$a_{14} \sqsubset \overline{a}_8^{111}$$

- en vertu de (3), on a

$$a_4^1 \sqsubset \overline{a}_3^2$$
$$a_{14} \sqsubset \overline{a}_3^1$$
$$a_4^2 \sqsubset \overline{a}_{15}^1$$

Le graphe de la figure 6.31 représente la relation $\lhd$ où les flèches droites représentent la relation $\ll$ et les flèches courbes représentent $\sqsubset$. Le graphe est acyclique et donc la formule 6.25 est valide.

Si on avait adopté une multiplicité $\mu(\overline{a}_3) = 1$, les noeuds $\overline{a}_3^1$ et $\overline{a}_3^2$ se seraient confondus, ainsi que $a_4^1$ et $a_4^2$, et le graphe aurait été cyclique voir figure 6.32. Or, quand le graphe est cyclique, on ne peut rien conclure.

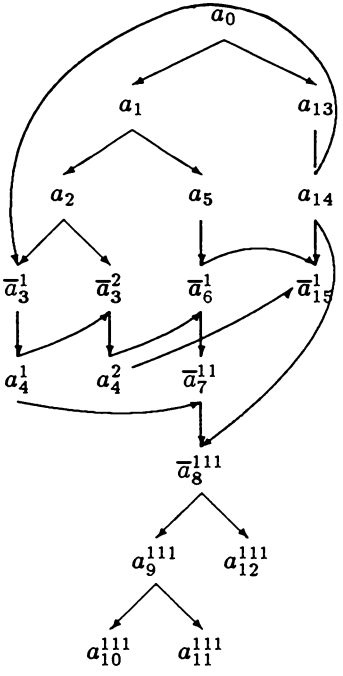



**Figure 6.31** *Graphe acyclique*

**Figure 6.32** *Graphe cyclique*

## 6.4. La méthode de résolution avec unification

### 6.4.1. *La démonstration automatique*

C'est en 1957 que paraît la première publication présentant un programme de démonstration de théorème par ordinateur. Le programme "The Logic Theory Machine" de Newell, Shaw et Simon était capable de prouver quelques théorèmes dans le calcul des propositions. Les auteurs précités faisaient appel à des méthodes heuristiques. Indépendamment l'un de l'autre, Paul Gilmore et Hao Wang ont cherché à étendre la démonstration automatique au calcul des prédicats, et à remplacer le recours aux méthodes heuristiques par l'emploi de procédures de preuves classiques en logique.

En 1960, Davis et Putnam présentent une procédure de preuve qui constitue un progrès majeur[8]. Elle s'appuyait sur le théorème de Herbrand, dont Bledsoe et Loveland résument ainsi la portée :

> La question de validité d'une formule du calcul des prédicats peut être ramenée à une série de questions à propos de la validité de formules propositionnelles, série en expansion continue (Bledsoe et Loveland, 1984, p. 7).

Comme une formule est valide si et seulement si sa négation est sémantiquement inconsistante, on peut reformuler ainsi le paragraphe précité : "la question de l'inconsistance d'une formule du calcul des prédicats peut être ramenée à une série de questions à propos de l'inconsistance de formules propositionnelles en expansion continue". C'est cette version du théorème de Herbrand qui est à la base de la procédure de preuve présentée en 1960 par Davis et Putnam. C'est elle aussi qui est utilisée par Quine dans sa "méthode des formes fonctionnelles" (Quine, 1973).

La démonstration automatique a vraiment démarré après la publication de l'article de J.A. Robinson : "A Machine-Oriented Logic Based on the Resolution Principle", en 1965. La logique de Robinson ne comporte aucun axiome et n'utilise qu'une seule règle d'inférence, la règle de résolution (cf. 4.4.3). A cela s'ajoute un algorithme pour faire les substitutions : l'*algorithme d'unification*.

[8] Elle posait cependant un problème important, auquel P. Prawitz s'est attaqué dans une contribution décisive en 1960.

La méthode de preuve de Robinson est à la base du langage Prolog inventé par Colmerauer et Roussel. Comme celle de Davis et Putnam et celle de Quine, elle repose sur le théorème de Herbrand.

Pour mesurer la portée exacte de l'apport décisif de Robinson, il est essentiel de distinguer la part que prennent dans sa logique trois ingrédients :

1. la réduction à la forme clausale,
2. la résolution,
3. l'unification.

### 6.4.2. *La réduction à la forme clausale*

Par une série de lemmes, nous allons démontrer que le problème de prouver l'inconsistance d'une formule du calcul des prédicats telle que, par exemple,

(6.26) $\neg(\forall x \forall y(Rxy \equiv Sx) \supset \forall y(Rxy \supset Sx))$,

se réduit à celui de prouver l'inconsistance d'un ensemble fini de formules de la forme suivante, que l'on appelle *clauses*, où les $l_i$ sont des littéraux positifs ou négatifs :

(6.27) $\{(l_1 \vee \cdots \vee l_m), \ldots, (l_1 \vee \cdots \vee l_n)\}$,

Après avoir énoncé et démontré les lemmes, nous appliquerons cette technique à la proposition 6.26, c'est-à-dire que nous réduirons la tâche de prouver l'inconsistance de 6.26 à celle de prouver l'inconsistance de l'ensemble formé par les quatre clauses ci-dessous dans lesquelles $x, x', x'', y, y', y'', v$ sont des variables individuelles, $S(-)$ et $R(-,-)$ des prédicats et $f(-,-)$ un symbole de fonction appelé foncteur de Skolem (cf. définition 5) :

(6.28) $\{\underbrace{\overline{R}xy \vee Sx}_{C_1}, \underbrace{\overline{S}x' \vee Rx'y'}_{C_2}, \underbrace{Rvf(x'',y'')}_{C_3}, \underbrace{\overline{S}v}_{C_4}\}$

### 6.4.3. *Définitions et lemmes justifiant la réduction*

**Définition 6.4 : clause.** On appelle clause une disjonction de littéraux, contenant ou non des variables libres d'individus.

Par exemple :

(6.29) $Fg(x) \vee \overline{H}ab \vee p$.

**Définition 6.5 : forme prénexe.** Une formule du calcul des prédicats — qu'elle soit ouverte ou fermée — est en forme prénexe si et seulement si tous ses quantificateurs sont en tête et ont une portée qui s'étend jusqu'à la fin de la formule.

Par exemple, la formule

(6.30) $\exists x \forall y((Fx \vee Gy) \supset (Kzx \equiv p))$

est en forme prénexe

**Définition 6.6 : préfixe.** On appelle préfixe la suite des quantificateurs dans une formule en forme prénexe.

Dans 6.30, le préfixe est

(6.31) $\exists x \forall y.$

**Définition 6.7 : matrice.** On appelle matrice la partie d'une formule en forme prénexe qui suit le préfixe. Par définition, cette partie de la formule ne contient pas de quantificateurs.

Dans 6.30, la matrice est

(6.32) $((Fx \vee Gy) \supset (Kzx \equiv p)).$

**Définition 6.8 : forme de Skolem.** Une formule du calcul des prédicats est en forme de Skolem si et seulement si

1. elle est en forme prénexe,
2. les quantificateurs existentiels qui venaient en tête dans le préfixe ont été remplacés par des constantes inédites $a_1, a_2, \ldots$, appelées constantes de Skolem (on appelle aussi ces constantes "constantes-témoins" (cf. lemme 9.11),
3. les quantificateurs existentiels qui étaient *précédés*, dans le préfixe, par des quantificateurs universels ont été supprimés, et les variables qu'ils liaient ont été remplacées par des symboles de fonction ne figurant pas déjà dans la formule, symboles de fonction appelés foncteurs de Skolem $(f(x_1, \ldots, x_m), g(y_1, \ldots, y_n), \ldots)$. Ces foncteurs prennent comme arguments les variables liées par les quantificateurs universels qui précédaient les quantificateurs existentiels supprimés.

Par exemple, la formule

(6.33) $\exists x \exists y \forall z \forall u \exists v((Rxy \supset (\overline{F}zuv \vee Kwa)) \wedge Gv),$

mise en forme de Skolem, devient

(6.34) $\forall z \forall u((Ra_1a_2 \supset (\overline{F}zuf(z,u) \vee Kwa)) \wedge Gf(z,u)),$

où $f$ est un foncteur de Skolem.

**Définition 6.9 : forme normale de Skolem.** Une formule du calcul des prédicats est en *forme normale* de Skolem si et seulement si

1. elle est en forme de Skolem (donc aussi en forme prénexe),
2. sa matrice est en forme conjonctive normale. En d'autres termes, sa matrice est une conjonction de clauses.

Par exemple, la formule 6.34, mise en forme normale, devient

(6.35) $\forall z \forall u(\underbrace{(\overline{R}a_1a_2 \vee \overline{F}zuf(z,u) \vee Kwa)}_{\text{clause 1}} \wedge \underbrace{Gf(z,u)}_{\text{clause 2}}).$

**Lemme 6.1.** Si $S'$ est une forme prénexe de $S$ (une même formule peut avoir plusieurs formes prénexes), alors $S'$ est logiquement équivalent à $S$.

*Démonstration.* Une démonstration détaillée se trouve dans Mendelson 1964, p.85–90. Nous nous bornerons à observer que pour transformer une formule $S$, qui n'est pas en forme prénexe, en une formule $S'$, qui est en forme prénexe, on utilise uniquement des équivalences telles que les lois de passage, les lois de conversion des quantificateurs ou les lois suivantes :

$$(6.36) \quad \forall x\Phi \wedge \forall x\Psi \equiv \forall x(\Phi \wedge \Psi),$$

$$(6.37) \quad \exists x\Phi \vee \exists x\Psi \equiv \exists x(\Phi \vee \Psi).$$

*Remarque.* Si dans $S$ des variables d'individu apparaissent à la fois à titre d'occurrence libre et d'occurrence liée, on met fin à cet état de choses en renommant les variables liées, ceci en vue de pouvoir appliquer les lois de passage.

Par exemple, si on a

$$(6.38) \quad \forall x Fx \supset (Gx \wedge \exists y Hy),$$

on commence par remplacer $x$ dans $\forall x Fx$ par $z$, et ensuite on applique les lois de passage de manière à obtenir une des formes prénexes de 6.38, par exemple :

$$(6.39) \quad \exists z \exists y(Fz \supset (Gx \wedge Hy)).$$

**Lemme 6.2.** Si $S'$ est une forme prénexe de $S$, alors $S'$ est inconsistant si et seulement si $S$ est inconsistant.

*Démonstration.* C'est une conséquence du fait que $S$ et $S'$ sont équivalents.

**Lemme 6.3.** Si $S'$ est une forme de Skolem de $S$, alors $S'$ est inconsistant si et seulement si $S$ est inconsistant.

*Démonstration.*

1. Si tous les existentiels du préfixe précèdent les universels, le lemme affirme que

$$(6.40) \quad \begin{array}{c} \exists x_1 \cdots \exists x_m \forall y_1 \cdots \forall y_n \Phi \text{ est inconsistant} \\ \text{si et seulement si} \\ \forall y_1 \cdots \forall y_n \Phi \text{ est inconsistant.} \end{array}$$

Bornons-nous au cas où $m = 1$, et prouvons que

$$(6.41) \quad \exists x\Phi \text{ est inconsistant si et seulement si } \Phi \text{ est inconsistant.}$$

On sait que

$$(6.42) \quad \models \forall x\Phi \text{ si et seulement si } \models \Phi,$$

c'est-à-dire, comme $\models \Phi$ si et seulement si $\neg\Phi$ est inconsistant,

(6.43) $\neg\forall x\Phi$ est inconsistant si et seulement si $\neg\Phi$ est inconsistant.

Par la loi des quantificateurs $\neg\forall x\Phi \equiv \exists x\neg\Phi$, on obtient

(6.44) $\exists x\neg\Phi$ est inconsistant si et seulement si $\neg\Phi$ est inconsistant.

On substitue $\neg\Phi$ à $\Phi$, on applique la loi de la double négation, et on obtient 6.41.

2. Si certains existentiels sont précédés d'universels, on peut, grâce à l'emploi des foncteurs de Skolem, revenir au cas où tous les existentiels précèdent les universels[9]. Quine résume ainsi le raisonnement :

> [...] considérons "$\forall x\exists yFxy$". Ce schéma nous dit que pour tout objet $x$ il existe un objet $y$, peut-être différent pour différents choix de $x$, et tel que $Fxy$. Généralement parlant, il affirme qu'il existe une *fonction* — une manière de choisir, pour chaque objet $x$, un objet dépendant $y_x$ [que Skolem noterait $f(x)$]. Le gain de cette formulation est que notre nouveau quantificateur existentiel tacite — "il existe une fonction" — vient avant "$(\forall x)$" et non après, comme le faisait "$(\exists y)$". Venant ainsi en premier, il peut rester non exprimé dans les preuves d'inconsistance (Quine, 1972, 1973, p. 198).

**Lemme 6.4.** Soient $C^1, \ldots, C^k$ des clauses au sens de la définition 6.4. Supposons qu'il y ait des variables communes à plusieurs clauses. Soient $C'^1, \ldots, C'^k$ les clauses obtenues en diversifiant les variables (libres) de manière à éviter cet état de choses. Par exemple, si $x$ figure dans $C^i$ et dans $C^j$, on remplacera $x$ par $x'$ dans $C^j$. L'ensemble des clauses $\{C'^1, \ldots, C'^k\}$ est inconsistant si et seulement si $\{C^1, \ldots, C^k\}$ est inconsistant.

*Démonstration.* Les variables libres des clauses sont prises dans le sens de la généralité (5.3.5). En d'autres termes, les quantificateurs sous-entendus sont des quantificateurs universels. D'autre part, l'ensemble fini de clauses est équivalent à une conjonction. Les formules

(6.45) $C^1 \wedge \ldots \wedge C^k$

et

(6.46) $C^{1'} \wedge \ldots \wedge C^{k'}$

sont équivalentes en raison de la loi de distribution du quantificateur universel sur la conjonction (cf. 5.3.9).

*Remarque.* Le renommage des variables libres n'est pas autorisé. Ce qui le justifie dans le cas des clauses, c'est précisément le fait que devant les variables qu'elles contiennent il y a un quantificateur universel sous-entendu.

---

[9] En logique du second ordre où on quantifie sur des variables de fonctions, on peut prouver $\forall x\exists y\Phi xy \equiv \exists f\forall x\Phi(x, f(x))$.

### 6.4.4. *Exemple de réduction*

Nous allons réduire la formule 6.26 de la section 6.4.2 :

$$\neg(\forall x \forall y(Rxy \equiv Sx) \supset \forall y(Rxy \supset Sx)).$$

**1ère phase :** renommage des variables liées. Ceci vise à éviter

1. que plusieurs quantificateurs combinés à une même variable apparaissent dans la formule. Ici, c'est le cas du quantificateur $\forall y$ qui figure deux fois dans la formule.
2. qu'une même variable apparaisse libre dans une certaine sous-formule et liée dans une autre :

(6.47) $$\neg(\forall v \forall w(Rvw \equiv Sv) \supset \forall y(Rxy \supset Sx)).$$

**2ème phase :** mise en forme prénexe.

(6.48) $$\forall v \forall w(Rvw \equiv Sv) \wedge \neg \forall y(Rxy \supset Sx),$$

(6.49) $$\forall v \forall w(Rvw \equiv Sv) \wedge \exists y \neg(Rxy \supset Sx),$$

(6.50) $$\forall v \forall w\,((Rvw \equiv Sv) \wedge \exists y \neg(Rxy \supset Sx)),$$

(6.51) $$\forall v \forall w \exists y\,((Rvw \equiv Sv) \wedge \neg(Rxy \supset Sx)).$$

**3ème phase :** mise en forme de Skolem.

(6.52) $$\forall v \forall w\,((Rvw \equiv Sv) \wedge \neg(Rxf(v,w) \supset Sx)).$$

**4ème phase :** mise en forme conjonctive normale ("normale fonctionnelle").

(6.53) $$\forall v \forall w\,((\overline{R}vw \vee Sv) \wedge (\overline{S}v \vee Rvw) \wedge Rxf(v,w) \wedge \overline{S}x).$$

**5ème phase :** distribution des quantificateurs et renommage.

(6.54) $$\forall v \forall w(\overline{R}vw \vee Sv) \wedge \forall v' \forall w'(\overline{S}v' \vee Rv'w') \wedge \forall v'' \forall w'' Rxf(v'',w'') \wedge \overline{S}x.$$

**6ème phase :** mise en forme clausale.

(6.55) $$\{\underbrace{(\overline{R}vw \vee Sv)}_{C_1}, \underbrace{(\overline{S}v' \vee Rv'w')}_{C_2}, \underbrace{Rxf(v'',w'')}_{C_3}, \underbrace{\overline{S}x}_{C_4}\}.$$

### 6.4.5. *Construction du lexique d'instantiation*

> Pour obtenir ce lexique, nous [...] repérons, pour commencer, les variables libres simples [...]. Elles font partie du lexique. Puis nous prenons les lettres de fonctions [...]; chaque lettre de fonction, combinée avec des membres quelconques du lexique utilisés comme indices, constitue à son tour un membre du lexique (Quine, 1972, 1973, p. 199).

Pour nous conformer à l'usage, nous y ajouterons les constantes de Skolem $a_1, \ldots, a_n$ issues de l'élimination des quantificateurs existentiels en tête. S'il n'y en avait aucun, on introduit la constante $a$.

Dans le cas de la formule 6.53, le lexique sera

$$\{a, f(a,a), f(f(a,a),a), f(a,f(a,a)), f(f(a,a),f(a,a)), \ldots\}. \tag{6.56}$$

**Définition 6.10 : clause fondamentale.** Appelons clauses fondamentales les clauses obtenues en substituant aux variables des clauses initiales (6.55) des termes du lexique (6.56), de toutes les manières possibles.

### 6.4.6. *Le théorème de Herbrand*

**Lemme 6.5.** Herbrand a établi que pour prouver l'inconsistance d'une expression en forme clausale, il suffisait d'inspecter les clauses fondamentales qui l'instancient. Dans *Méthodes de logique*, Quine formule ce lemme de manière concise :

> Il n'est jamais nécessaire, s'agissant d'établir l'inconsistance, de produire des instantiations avec autre chose que le lexique (Quine, 1972, 1973, p. 199).

*Démonstration.* Ce lemme-ci, comme le théorème de Herbrand qui suit, sera démontré à la section 8.3.7. Quand le lexique contient un foncteur, il est infini et l'ensemble des clauses fondamentales qui instancient la formule dont on veut prouver l'inconsistance l'est aussi, mais Herbrand démontre qu'un sous-ensemble fini de telles clauses suffit.

**Théorème 6.6.** Un ensemble $S$ de clauses est sémantiquement inconsistant si et seulement s'il existe un ensemble de clauses fondamentales $S'$, inconsistant et fini.

Le théorème de Herbrand recèle une méthode constructive pour trouver un tel ensemble : on énumère *sans répétition* et *sans omission* les membres de l'ensemble infini des instances fondamentales des clauses, en prenant soin de s'arrêter *périodiquement* pour tester la consistance de l'ensemble fini de propositions déjà obtenues.

Dans le cas des clauses fondamentales énumérées en 6.55, on obtient aisément une inconsistance évidente (deux littéraux complémentaires) en procédant aux substitutions

$$\begin{aligned} \sigma_1 &: \langle v, a\rangle, \langle w, f(a,a)\rangle, \\ \sigma_2 &: \langle v', a\rangle, \langle w', f(a,a)\rangle, \\ \sigma_3 &: \langle v'', a\rangle, \langle w'', a\rangle. \end{aligned} \tag{6.57}$$

On obtient alors l'ensemble

(6.58) $\{\underbrace{(\overline{R}af(a,a) \vee Sa)}_{C_1}, \underbrace{(\overline{S}a \vee Raf(a,a))}_{C_2}, \underbrace{Raf(a,a)}_{C_3}, \underbrace{\overline{S}a}_{C_4}\}$.

En appliquant le syllogisme disjonctif sur $C_1$ et $C_3$, on obtient $Sa$ qui contredit $C_4$.

*Remarque.* Si on n'avait pas opéré le renommage dans 6.54, la clause $C_3$ aurait été

(6.59) $Rxf(v,w)$,

et on n'aurait pas pu obtenir deux littéraux complémentaires ($\overline{R}af(a,a)$ et $Raf(a,a)$) en substituant dans 6.55. La diversification des variables — le renommage — est autorisée par la distribution de $\forall$ sur $\wedge$; elle n'est pas seulement légitime, elle est également souvent indispensable à la production d'une inconsistance (cf. section 6.4.9).

### 6.4.7. *La méthode de résolution*

Puisque les clauses fondamentales appartiennent au calcul des propositions, on peut utiliser la règle de résolution pour faire ce test. On commence par former un ensemble fini de clauses fondamentales, en instantiant les clauses initiales à l'aide de termes puisés dans le lexique. Appelons $E(S)$ le résultat de cette instantiation. On applique alors à $E(S)$ la règle de résolution décrite en 4.4.3, et on ajoute à l'ensemble initial $E(S)$ un ou plusieurs résolvants issus de cette application[10]. Appelons $R(E(S))$ l'ensemble ainsi obtenu. On applique alors la résolution à ce nouvel ensemble, de manière à obtenir $R^2(E(S))$. Puisque chaque ensemble est obtenu par l'union de résolvants à l'ensemble précédent, chaque ensemble est contenu dans son successeur. On a

(6.60) $E(S) \subseteq R(E(S)) \subseteq R^2(E(S)) \ldots \subseteq R^n(E(S))$.

On continue ainsi indéfiniment. On s'arrête si on arrive à un $R^i(E(S))$ dans lequel figurent deux littéraux complémentaires, c'est-à-dire une inconsistance flagrante. Cette procédure est non déterministe.

Soit, par exemple, l'ensemble

(6.61) $\{Px \vee Qy, Py \vee \overline{Q}x, \overline{P}x \vee Qx, \overline{P}y \vee \overline{Q}y\}$.

On commence par définir le domaine de Herbrand de cet ensemble. Comme il n'y a pas de foncteur, ce domaine est $\{a\}$. On procède ensuite à l'instantiation de $S$ en puisant dans le domaine de Herbrand de $S$, c'est-à-dire $H_S$. On obtient alors l'ensemble de clauses fondamentales

[10] Rappelons que l'on appelle résolvant des clauses $C_1$ et $C_2$ la clause $C_3$ obtenue en formant la disjonction des membres de $C_1$ et de $C_2$ qui subsistent après l'effacement des littéraux complémentaires $\ell$ et $\overline{\ell}$ qui figuraient dans $C_1$ et $C_2$.

(6.62) $E(S) = \{Pa \vee Qa, Pa \vee \overline{Q}a, \overline{P}a \vee Qa, \overline{P}a \vee \overline{Q}a\}.$

Le résolvant des deux premières clauses de $E(S)$ est $Pa$. Ajoutons-le à $E(S)$. Nous obtenons, à l'issue de la première résolution sur *clauses fondamentales* de $S$, le nouvel ensemble :

(6.63) $R(E(S)) = E(S) \cup \{Pa\}.$

Comme

(6.64) $R(E(S)) = \{Pa \vee Qa, Pa \vee \overline{Q}a, \overline{P}a \vee Qa, \overline{P}a \vee \overline{Q}a, Pa\},$

une nouvelle résolution est possible.

Le résolvant des deux dernières clauses de $R(E(S))$ est $\overline{Q}a$. Ajoutons-le à $R(E(S))$, nous obtenons la deuxième résolution sur clauses fondamentales de $S$ :

(6.65) $R^2(E(S)) = R(E(S)) \cup \{\overline{Q}a\}.$

Comme

(6.66) $R^2(E(S)) = \{Pa \vee Qa, Pa \vee \overline{Q}a, \overline{P}a \vee Qa, \overline{P}a \vee \overline{Q}a, Pa, \overline{Q}a\},$

une nouvelle résolution est possible.

Le résolvant de la troisième et de la cinquième clause est $\overline{P}a$. Ajoutons-le à $R^2(E(S))$, nous obtenons la troisième résolution sur clauses fondamentales de $S$ :

(6.67) $R^3(E(S)) = R^2(E(S)) \cup \{\overline{P}a\} =$

(6.68) $\{Pa \vee Qa, Pa \vee \overline{Q}a, \overline{P}a \vee Qa, \overline{P}a \vee \overline{Q}a, \underline{Pa}, \overline{Q}a, \underline{\overline{P}a}\}.$

Cet ensemble contient deux littéraux complémentaires. Dès lors, il est manifestement inconsistant.

Dans la pratique, cette méthode est fort lourde. Robinson ne l'a introduite qu'à des fins théoriques. On va voir en effet qu'en inversant l'instantiation et la résolution, on obtient une méthode beaucoup plus performante.

### 6.4.8. *L'unification*

Dans la méthode de preuve d'inconsistance qui repose sur le théorème de Herbrand, on commence par *instancier* — ce que Robinson appelle "saturer" — les clauses en puisant dans l'univers de Herbrand, et ensuite on applique la résolution. J.A. Robinson a démontré en 1965 que la résolution et la saturation (instantiation) sont semi-commutatives. En d'autres termes, on a

(6.69) $R(E(S)) \subseteq E(R(S)),$

et, sous une forme plus générale,

(6.70) $R^n(E(S)) \subseteq E(R^n(S)).$

La portée de cette semi-commutativité est clairement indiquée dans le passage suivant :

> [...] N'importe quelle clause fondamentale qui peut être obtenue en instantiant *d'abord* une paire de clauses $C$, $D$ de l'ensemble $S$ dans un sous-ensemble quelconque de l'univers de Herbrand et en formant *ensuite* un résolvant fondamental des deux instances résultantes peut également être obtenue en instantiant sur $P$ l'un des résolvants finiment nombreux de $C$ et $D$ (Robinson, 1965, 1983, p. 403).

En d'autres termes, on peut *intervertir* l'instantiation et la résolution : au lieu de commencer par instancier les clauses et de leur appliquer la résolution, on commencera par appliquer la résolution aux clauses et ensuite on les instanciera. On voit tout de suite l'intérêt qu'il y a à permuter les deux opérations : les résolutions effectuables sur une paire de clauses sont en nombre fini — les clauses étant finies — tandis que le nombre d'instances possible d'une paire de clauses dans l'univers de Herbrand est infini dès que l'univers de Herbrand est infini.

Comment faire pour appliquer la résolution avant d'instancier ? Pour voir le problème, considérons les clauses $Fxy \vee Gy$ et $Kv \vee \overline{G}z$. Pour que $Gy$ et $\overline{G}z$ s'annulent, il faut d'abord effacer la différence entre $y$ et $z$ par une même constante puisée dans l'univers de Herbrand, mais il y a un moyen plus rapide d'obtenir ce résultat : c'est de travailler directement sur les clauses elles-mêmes sans devoir considérer leurs instances fondamentales. L'idée vient de Robinson : il s'agit de chercher une substitution qui, appliquée à deux littéraux, permet de les transformer en deux instances du même littéral. On appelle *unificateur* une substitution qui a ce pouvoir. Parmi les substitutions qui permettent d'unifier deux littéraux, il peut en exister qui sont *plus générales* que d'autres. Ces dernières ne sont pas nécessairement uniques. Nous les appellerons *unificateurs principaux*. Plus précisément, un unificateur est dit principal si pour tout unificateur $v$, il existe une substitution $w$ telle que $v = w \circ u$[11].

Par exemple, en substituant $z$ à $y$ dans les littéraux complémentaires $Gy$ et $\overline{G}z$, on les "unifie" et on obtient :

(6.71) $\underbrace{Fxz \vee Gz}_{C_1}, \underbrace{Kv \vee \overline{G}z}_{C_2}$

ce qui permet, ultérieurement, d'appliquer la règle de résolution et d'obtenir le résolvant

(6.72) $Fxz \vee Kv.$

Le but de l'unification (substitution identifiant des littéraux au signe près) est d'ailleurs de rendre possible l'application de la règle de résolution.

[11] Soient $Gxz$ et $\neg Gya$. Par $v = \langle x, a\rangle, \langle y, a\rangle, \langle z, a\rangle$, on obtient $Gaa$ et $\neg Gaa$. Si, au contraire, on prend $u = \langle x, y\rangle$ et $w = \langle z, a\rangle$, et qu'on les compose, on obtient $Gya$ et $\neg Gya$ dont les deux littéraux précédents sont des instances.

Si la variable d'un littéral $C_i$ d'une clause $C_j$ apparait également dans d'autres littéraux de la même clause, elle subit aussi l'effet de la substitution. Par exemple, la variable $y$ qui figurait dans le littéral $Gy$ qu'on cherche à unifier avec $\overline{G}z$ apparait aussi dans $Fxy$. Dès lors, la substitution l'affecte aussi.

L'instantiation dans l'univers de Herbrand est une substitution *exogène* : elle puise ses instances dans un univers dont les membres sont généralement extérieurs à la formule. Ces membres sont en nombre infini dès que l'une des clauses contient un foncteur. Au contraire, l'unification est une substitution *endogène* : elle puise ses instances exclusivement dans les symboles des paires initiales, ou des paires dérivées de celles-ci par un nombre fini d'applications du principe de substitution. En intervertissant résolution et instantiation dans l'univers de Herbrand, on rend le recours à l'univers de Herbrand superflu. L'unification suffit.

### 6.4.9. *Contraintes sur l'unification.*

Soit la paire de littéraux suivante $A_1, A_2$. On déplace un pointeur $\uparrow$ qui indique quels arguments du prédicat présentent une différence. Ici, les trois arguments le font :

(6.73) $\quad P(\underset{\uparrow}{f(z)}, z, g(b,u)), P(\underset{\uparrow}{f(a)}, x, g(u,v)).$

On commence par former la *paire de discordance* $D_0 = \{a, z\}$. La lettre $a$ étant une constante et $z$ une variable, la substitution unifiante sera $\langle z, a\rangle$, ce qui se lit : "$z$ devient $a$". Notons que l'exigence d'uniformité nous oblige à remplacer aussi la deuxième occurrence de $z$ dans $A_2$ par $a$.

L'unification n'est possible que si l'ensemble de discordance contient comme membres quelque variable $v_K$ et quelque terme $t_K$. S'il n'y avait que des constantes, aucune substitution, donc aucune unification ne serait possible. Par exemple, on ne peut unifier $Pa, Pf(x)$. En effet, $a$ et $f$ sont des constantes. En revanche, on peut unifier $Pf(y), Pf(x)$ car les *sous-termes* $x$ et $y$ sont des variables. Une deuxième restriction doit être notée : le terme $t_K$ ne peut pas contenir $v_K$. Si $t_K$ contenait $v_K$, toute substitution déclencherait une régression à l'infini. Par exemple,

(6.74) $\quad Px, Pf(x)$

n'est pas unifiable. En effet, $v_K = x$ et $t_K = f(x)$, donc $t_K$ contient $v_K$. Si on substituait $f(x)$, c'est-à-dire $t_K$, à $x$, c'est-à-dire à $v_K$, on obtiendrait

(6.75) $\quad Pf(x), Pf(f(x)),$

puis

(6.76) $\quad Pf(f(x)), Pf(f(f(x))),$

et ainsi de suite.

La diversification des variables d'une clause à l'autre diminue le risque de blocage survenant lorsque le terme $t_K$, candidat au rôle de substitut de la variable $x$, contient lui-même cette variable (cf. 6.4.6, remarque).

### 6.4.10. *Algorithmes d'unification et de résolution*

On écrit sur une même ligne les deux littéraux que l'on veut unifier et on place le pointeur sous le premier élément de chacun d'eux (terme, sous-terme, ...) par lequel ils diffèrent. On efface les différences par des substitutions successives.

Il existe un algorithme pour découvrir les unificateurs principaux quand ils existent. Soit $\mathcal{S}$ un ensemble de clauses dont on cherche à prouver l'inconsistance. L'algorithme peut être décrit en ces termes :

Tant que les deux littéraux $(\ell, \ell')$ que l'on cherche à unifier ne sont pas identiques

* Placer le pointeur sous le premier terme de $\ell$ et de $\ell'$ (sous-terme, sous-sous-terme, ...) par lequel ils diffèrent et former la première paire de discordance.

- Si les membres de cette paire sont des constantes, s'arrêter et déclarer les littéraux *non unifiables*.
- Si l'un des membres de cette paire est un terme $t_k$ et l'autre une variable $v_k$ et si $t_k$ contient $v_k$, s'arrêter et déclarer les littéraux *non unifiables*.
- Sinon substituer $t_k$ à $v_k$ (si $t_k$ est une variable, on peut soit substituer $t_k$ à $v_k$, soit substituer $v_k$ à $t_k$. L'algorithme n'est donc pas déterministe) et passer à *.

**Figure 6.33** *Algorithme d'unification*

Les deux pointeurs se déplacent simultanément vers la droite. Comme l'écrit Loveland :

> ...les portions à gauche des pointeurs sont toujours identiques. Le processus ressemble au mouvement d'une fermeture éclair se déplaçant de gauche à droite, les expressions à droite des pointeurs pouvant différer, mais celles qui sont à gauche ayant été rendues semblables (Loveland, 1978, p. 76).

Soient, par exemple[12], les deux littéraux suivants $\ell$ et $\ell'$, formé d'un prédicat ternaire $R$ (par exemple "- est entre - et -") et de trois termes ("$S$" est le symbole de fonction "successeur de") :

(6.77) $R(5, x, !(S(y)))$ et $R(z, +(z, w), !(w))$.

La première paire de discordance est $D_1 = \{5, z\}$. La seule substitution unifiante possible est $\sigma_1 = \langle z, 5\rangle$ qui se lit "$z$ devient 5". Après l'exécution de cette substitution, on déplace le pointeur vers la droite jusqu'au prochain terme par lequel les littéraux $\ell$ et $\ell'$ diffèrent

(6.78) $R(5, x, !(S(y)))$ et $R(5, +(5, w), !(w))$.
$\uparrow$ $\uparrow$

La deuxième paire de discordance est $D_2 = \{x, +(5, w)\}$. La seule substitution unifiante possible est $\sigma_2 = \langle x, +(5, w)\rangle$. Après l'exécution de cette substitution, on déplace le pointeur vers la droite jusqu'au prochain terme différent. Cette fois, il s'agit d'un sous-terme :

(6.79) $R(5, +(5, w), !(S(y)))$ et $R(5, +(5, w), !(w))$.
$\uparrow$ $\uparrow$

La troisième paire de discordance est $D_3 = \{w, S(y)\}$. La seule substitution possible est $\sigma_3 = \langle w, S(y)\rangle$. Après exécution, les deux littéraux sont semblables. Ils sont devenus deux occurrences du même littéral (dans la terminologie de Pierce, deux "échantillons" (*tokens*) du même "type"; cf. 1.5.1) :

(6.80) $R(5, +(5, S(y)), !(S(y)))$ et $R(5, +(5, S(y)), !(S(y)))$.

La substitution est une fonction. A ce titre, on peut *composer* une substitution[13] avec une autre. Ce que l'algorithme d'unification calcule pour nous, c'est précisément le résultat de compositions successives de fonctions. Sans son aide, ce calcul aurait été laborieux.

Soient les fonctions

(6.81) $\sigma_1 = \langle z, 5\rangle$ et $\sigma_2 = \langle x, +(5, w)\rangle$.

La composition de ces deux fonctions donne comme résultat

(6.82) $\sigma_3 = \sigma_2 \circ \sigma_1 = \{\langle z, 5\rangle, \langle x, +(5, w)\rangle\}$.

On a ensuite $\sigma_4 = \langle w, S(y)\rangle$. La composition de $\sigma_3$ avec $\sigma_4$ donne

(6.83) $\sigma_5 = \sigma_4 \circ \sigma_3 = \{\langle z, 5\rangle, \langle x, +(5, S(y))\rangle, \langle w, S(y)\rangle\}$.

*Remarque.* Quand on compose deux substitutions telles que le résultat de la première est une expression comportant une variable libre qui peut servir d'argument à la deuxième substitution, il ne suffit pas d'effectuer la deuxième sur le résultat de la première, il faut également conserver la deuxième. Par

[12] Cet exemple est adapté de Manna 1974.

[13] Etant donné les fonctions $f : X \mapsto Y$ et $g : Y \mapsto Z$, il existe une fonction $h : X \mapsto Z$ telle que $h(x) = g(f(x))$, ce qui s'écrit aussi $h(x) = (g \circ f)(x)$ (Rasiowa, 1973, p. 45).

exemple, la composition de $\sigma_1 = \langle x, f(y)\rangle$ et de $\sigma_2 = \langle y, g(z)\rangle$ n'est pas $\sigma_3 = \langle x, f(g(z))\rangle$ mais est $\sigma_3 = \{\langle x, f(g(z))\rangle, \langle y, g(z)\rangle\}$. Pour la même raison, la composition de $\langle x, y\rangle$ et $\langle y, x\rangle$ donne $\langle y, x\rangle$. En effet, on n'écrit pas la substitution nulle $\langle x, x\rangle$.

La quête d'un unificateur principal qui unifie deux littéraux à l'aide de l'algorithme d'unification intervient *notamment* lors de l'application de la première instruction de l'algorithme de résolution ci-dessous.

Tant que $\mathcal{F} \notin \mathcal{S}$,
    choisir $l_1, l_2, C_1, C_2$ tels que :
        $C_1$ et $C_2$ sont des clauses ou facteurs de clauses;
        $l_1 \in C_1$, $\neg l_2 \in C_2$ et $l_1, l_2$ sont unifiables;
    calculer la clause résolvante $r$ obtenue au moyen d'un unificateur principal;
    remplacer $\mathcal{S}$ par $\mathcal{S} \cup \{r\}$.

**Figure 6.34** *Algorithme de résolution*

### 6.4.11. *Application*

Démontrons une dernière fois l'inconsistance de l'ensemble de formules

(6.84) $\{\forall x\forall y\forall z((Fxy \wedge Fyz) \supset Gxz), \forall x\exists y Fyx, \neg\forall x\exists y Gyx\}$.

La mise en forme clausale produit trois clauses :

(6.85) I : $\overline{F}xy \vee \overline{F}yz \vee Gxz$;

(6.86) II : $Fg(x')x'$;

(6.87) III : $\overline{G}y'a$.

On cherche à unifier les deux littéraux :

(6.88) $Gxz$ et $\overline{G}y'a$.
$\uparrow \quad \uparrow$

La première paire de discordance est $D_1 = \{x, y'\}$. Une manière d'éliminer la discordance est d'adopter la substitution $\sigma_1 = \langle x, y'\rangle$ qui remplace $x$ par $y'$. Cette substitution se répercute dans la clause I dont $Gxz$ fait partie. Cette clause devient :

(6.89) I' : $\overline{F}y'y \vee \overline{F}yz \vee Gy'z$.

Les deux littéraux initiaux sont devenus

(6.90) $Gy'z$ et $\overline{G}y'a$.
$\uparrow \quad \uparrow$

La paire de discordance est à présent $D_2 = \{z, a\}$. La seule manière de l'éliminer est d'adopter la substitution $\sigma_2 = \langle z, a\rangle$.

La clause I' devient

(6.91) I" : $\overline{F}y'y \vee \overline{F}ya \vee Gy'a.$

Par résolution, sur I" et III, on obtient

(6.92) IV : $\overline{F}y'y \vee \overline{F}ya.$

Cherchons à unifier le littéral $\overline{F}ya$ avec II, c'est-à-dire

(6.93) $\overline{F}ya$ et $Fg(x')x'$
$\uparrow \quad \uparrow$

La première paire de discordance est $D_1 = \{y, g(x')\}$. La seule substitution possible est $\sigma_3 = \langle y, g(x')\rangle$. On obtient

(6.94) IV' : $\overline{F}y'g(x') \vee \overline{F}g(x')a.$

Les deux littéraux initiaux ($\overline{F}ya$ et II) sont devenus $\overline{F}g(x')a$ et $Fg(x')x'$. La deuxième paire de discordance est $D_2 = \{x', a\}$. La seule substitution possible est $\sigma_4 = \langle x', a\rangle$. On obtient

(6.95) IV" : $\overline{F}y'g(a) \vee \overline{F}g(a)a$ et

(6.96) II' : $Fg(a)a.$

Par résolution sur IV" et II', on obtient

(6.97) V : $\overline{F}y'g(a)$

que l'on cherche à unifier avec II :

(6.98) $F\underbrace{g(x')}_{\uparrow}x'$ et $\overline{F}\underbrace{y'}_{\uparrow}g(a)$

La première paire de discordance est $D_1 = \{g(x'), y'\}$. La seule substitution possible est $\sigma_5 = \langle y', g(x')\rangle$. On obtient

(6.99) V' : $\overline{F}g(x')g(a)$

Dans $\overline{F}g(x')g(a)$ et $Fg(x')x'$ la paire de discordance est $D_2 = \{g(a), x'\}$ et
$\uparrow \quad \uparrow$
la seule substitution possible est $\sigma_6 = \langle x', g(a)\rangle$. On l'effectue et on obtient ainsi finalement

(6.100) $\overline{F}g(g(a))g(a)$ , $Fg(g(a))g(a)$

c'est-à-dire $\mathcal{F}$.

### 6.4.12. *Le fondement de la méthode de Robinson*

La semi-commutativité entre les opérations d'instantiation et de résolution conduit au lemme du relèvement (*lifting lemma*) qui s'énonce ainsi :

(6.101) Si $C_1'$ est une instance de $C_1$ et $C_2'$ une instance de $C_2$ et si $C'$ est le résolvant de $C_1'$ et de $C_2'$, alors il existe un résolvant $C$ de $C_1$ et de $C_2$ tel que $C'$ est une instance de $C$.

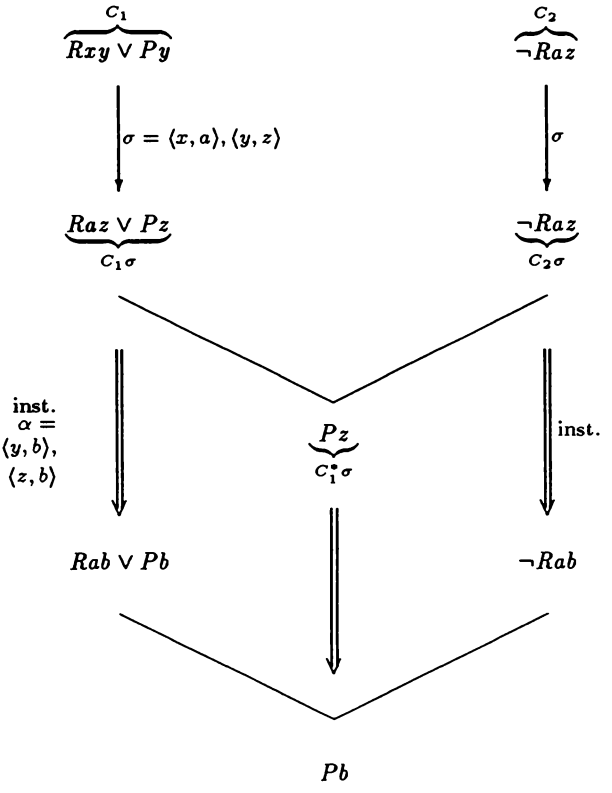



**Figure 6.35** *Graphe du relèvement*

*Exemple.* On voit sur la figure 6.35 que la suite d'opérations "*unifier* $C_1$ et $C_2$ grâce à la substitution $\sigma$, *résoudre* en $R$, et *instancier* le résultat de cette résolution" a le même effet que la suite "*unifier* $C_1$ et $C_2$ grâce à la substitution $\sigma$, *instancier* $C_1\sigma$, $C_2\sigma$, *résoudre*". Schématiquement, $\langle u, r, i \rangle$ équivaut à $\langle u, i, r \rangle$ (cf. figure 6.36).

Grâce au *lifting lemma*, on peut démontrer le thèorème du relèvement (*lifting theorem*) qui permet de *hisser* une démonstration du calcul des propositions (faite de clauses fondamentales) au niveau du calcul des prédicats (faite de clauses ouvertes).

### 6.4.13. *Les limites de la méthode de résolution*

La *méthode de résolution* a été perfectionnée. Elle a été étendue à la logique modale (Fariñas). Certaines logiques sont réfractaires à la mise en

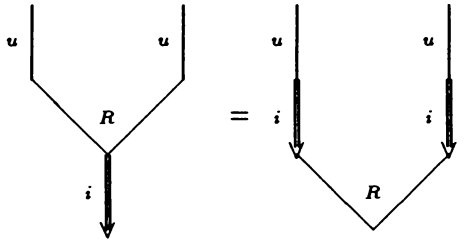



**Figure 6.36** *Schéma du relèvement*

forme clausale. C'est pour contourner cet obstacle que L. Wallen a développé un autre style de démonstration automatique qui utilise cependant le précieux *algorithme d'unification*. Notons qu'on peut éviter la mise sous forme clausale.

*Rappel.* Pour que la résolution binaire jouisse de la propriété de complétude, il faut y ajouter la règle de *factorisation* (unification au sein d'une même clause, puis simplification sur base de l'idempotence).

### 6.4.14. *Unification et méthode axiomatique*

L'algorithme d'unification ne doit pas être ravalé au rang d'une "sous-routine" de l'algorithme de résolution. Il peut se révéler utile dans le cadre d'autres méthodes analytiques comme, par exemple, la méthode des matrices de Wallen (cf. 6.3.10). Il peut même être utilisé dans le cadre de méthodes synthétiques comme la méthode axiomatique.

Jacques Pitrat qui a découvert l'algorithme d'unification en 1966 indépendamment de John Robinson a élaboré un démonstrateur automatique de théorèmes qui fait intervenir cet algorithme comme instrument de recherche des substitutions optimales dans un système axiomatique.

Quand, en 3.4.7, nous cherchions à démontrer le théorème 3.2

(6.102) $p \supset (p \supset (p \supset p))$

nous sommes partis du théorème 3.1

(6.103) $((q \supset p) \supset r) \supset (p \supset r)$

et de l'axiome **A.3** où nous marquerons les variables propositionnelles à l'aide de primes. On a en effet intérêt à *diversifier* les variables des formules avant l'unification (cf. 6.4.9 et on a le droit de le faire (lemme 6.4).

(6.104) $p' \supset (q' \supset p')$

Nous avons cherché des substitutions par lesquelles la transformée de l'axiome **A.3** devienne semblable à l'antécédent de la transformée du théorème 3.1 de manière à pouvoir appliquer la règle de détachement. Ce que nous avons fait

par tâtonnements, nous pouvons le faire à l'aide de l'algorithme d'unification moyennant quelques aménagements à la règle de substitution (on travaille sur l'arbre syntaxique de la formule plutôt que sur la formule elle-même. Pour une présentation détaillée, voir Laurière 1987, p.99-101). Il s'agit en effet d'*unifier* **A.3** à l'antécédent de 3.1. L'application de l'algorithme d'unification découvre les deux substitutions suivantes :

$$\sigma_1 = \langle r, q' \supset (q \supset p) \rangle$$
$$\sigma_2 = \langle p', q \supset p \rangle$$

Grâce à ces deux substitutions, nous obtenons les transformées suivantes du théorème 3.1 et de l'axiome **A.3**

(6.105) $((q \supset p) \supset (q' \supset (q \supset p))) \supset (p \supset (q' \supset (q \supset p)))$

(6.106) $(q \supset p) \supset (q' \supset (q \supset p))$

d'où nous obtenons, par détachement, le théorème

(6.107) $p \supset (q' \supset (q \supset p))$

ou, après élimination des primes

(6.108) $p \supset (r \supset (q \supset p))$

On observera que ce théorème est plus général que le théorème 3.2 qu'il compte parmi ses instances.

(6.109) $p \supset (p \supset (p \supset p))$

On voit à présent toute la richesse de la notion d'unificteur principal qui est à la base de l'algorithme d'unification : non seulement cet algorithme nous permet de trouver mécaniquement ce que nous n'avions d'abord trouvé que par tâtonnements, mais plus fondamentalement il nous permet de découvrir des *résultats plus généraux*.

## 6.5. Pour clore

Dans *Questions de Psychologie* (1988), F. Pire observe que si l'homme est un animal rationnel selon la définition aristotélicienne bien connue et même "le seul système qui allie le sens de la réalité au traitement formel des données", il n'en demeure pas moins que "sa rationalité est souvent un leurre".

Les psychologues du raisonnement ont mis en lumière la fréquence des fautes de logique dans le raisonnement quotidien le plus élémentaire. F. Pire rapporte dans l'ouvrage précité l'expérience suivante réalisée par Johnson-Laird et P. Watson :

> La tâche paraît simple. L'expérimentateur dispose devant le sujet quatre cartes portant les signes suivants :
>
> *E* *K* 4 7

> Le sujet sait que chaque carte a d'un côté un nombre, de l'autre côté une lettre. On lui dit que si une carte a une voyelle d'un côté, elle a de l'autre côté un nombre pair. La tâche consiste à décider quelles cartes il suffit de soulever pour savoir si la règle est respectée. Beaucoup de sujets comprendront vite qu'il suffit de lever deux cartes, mais la plupart font un mauvais choix, levant le $E$ et le 4 quand il s'agit de lever le $E$ et le 7 (Pire, 1988, p. 99).

Commençons par formaliser le raisonnement correct qui permet de résoudre le problème, c'est-à-dire de désigner les cartes qu'il faut retourner pour savoir si la règle est respectée.

Soient

- $V(-)$ le prédicat à une place "– est une voyelle",
- $P(-)$ le prédicat à une place "– désigne un nombre pair",
- $f(-)$ le foncteur à une place "le signe au verso de la carte portant le signe –".

La règle s'énonce :

> Pour tout $x$, si $x$ est une voyelle, le signe au verso de la carte portant l'inscription $x$ désigne un nombre pair,

ce qui s'écrit :

(6.110) $\forall x(Vx \supset Pf(x))$.

On admet *tacitement* les quatre vérités de fait :

(6.111) $V$"$E$",

(6.112) $\neg V$"$K$",

(6.113) $P$"4",

(6.114) $\neg P$"7",

et la loi suivante :

(6.115) *Le verso du verso redonne le recto*,

qui s'écrit formellement

(6.116) $\forall x(x = f(f(x)))$.

Ce qui est vrai pour tout $x$, l'est en particulier pour $E$. On peut donc écrire

(6.117) $V$"$E$" $\supset Pf($"$E$"$)$.

De la règle 6.110, en appliquant la loi logique de contraposition "Tous les $\Phi$ sont $\Psi$ si et seulement si tous les non-$\Psi$ sont non-$\Phi$" :

(6.118) $\forall x(\Phi \supset \Psi) \equiv \forall x(\neg\Psi \supset \neg\Phi)$,

on obtient

(6.119) $\forall x(\neg Pf(x) \supset \neg Vx)$,

et en particulier, par substitution de $f(x)$ à $x$,

(6.120) $\forall x(\neg P f(f(x)) \supset \neg V f(x))$.

Appliquant une instance de l'égalité 6.116 dans 6.120, on obtient

(6.121) $\forall x(\neg P x \supset \neg V f(x))$,

et en particulier, par substitution de "7" à $x$,

(6.122) $\neg P$"7" $\supset \neg V f$("7").

Le *modus ponens* sur $V$"$E$" $\supset P f$("$E$"), qui est une instance de 6.110, et $\neg P$"7" ainsi que sur $\neg P$"7" $\supset \neg V f$("7") et $V$"$E$" permet de faire une prédiction que l'on peut vérifier en retournant les cartes portant les inscriptions "$E$" et "7". Si les deux prédictions se vérifient, la règle 6.110 est respectée. Sinon, elle ne l'est pas.

Les sujets qui font le mauvais choix et lèvent le 4 font en réalité le raisonnement suivant :

> Une voyelle au verso entraîne un nombre pair au recto, donc un nombre pair au recto entraîne une voyelle au verso.

Ce raisonnement recèle un pas fallacieux, à savoir l'utilisation de la formule non valide

(6.123) $\forall x(\Phi \supset \Psi) \equiv \forall x(\Psi \supset \Phi)$,

ou du schéma d'inférence qui lui correspond.

# Chapitre 7
# La sémantique du calcul des prédicats

## 7.1. Deux sortes de formalismes

En 2.1.8, nous avons présenté l'opposition entre la *logique comme langage* et la *logique comme calcul*, opposition dont J. van Heijenoort et J. Hintikka ont montré l'importance. Au moment d'aborder la sémantique du calcul des prédicats grâce à laquelle ce calcul sera transformé en langage, il est bon de rappeler succintement en quoi l'idéographie de Frege et le formalisme de Peano ou de Russell diffèrent du calcul des prédicats.

Comme l'écrit Hodges,

> [les] langages de Frege et de Peano sont réellement des fragments du langage naturel, dont la signification et la grammaire ont été rendus plus rigoureuses et avec des symboles particuliers à la place des mots conventionnels (Hodges, 1975–1976, p. 143).

Le *calcul des prédicats* appartient à un autre type de langage formalisé. Il est un langage formalisé dépourvu de toute interprétation a priori et auquel une interprétation initiale doit être ajoutée de l'extérieur. La dissociation entre le langage et son interprétation n'est plus opérée ici dans le but de constituer une sorte de champ stérile — le formalisme pur — dans lequel les démonstrations sont ramenées à des calculs. Elle est opérée, au contraire, pour permettre des interprétations variées.

## 7.2. Les ingrédients d'une sémantique

Les langues naturelles et les langues formelles possèdent en commun deux caractéristiques, dont une sémantique — c'est-à-dire une théorie de la signification — doit pouvoir rendre compte.

Connaître une langue, c'est pouvoir comprendre ou produire des phrases *inédites* (comme l'est sans doute la phrase "En ce premier janvier de l'an trois mille, le dernier théorème de Fermat a été démontré"). Cet aspect de la compétence sémantique des sujets parlants repose sur une caractéristique de la langue que certains linguistes appellent "productivité". Cette première caractéristique a été bien mise en lumière par Wittgenstein dans le *Tractatus logico-philosophicus* :

> Il appartient à l'essence d'une phrase (*Satz*) qu'elle puisse nous communiquer un sens nouveau (Wittgenstein, 1921, 4.27),

et

> Une phrase (nouvelle) doit utiliser des expressions anciennes pour nous communiquer un nouveau sens (Wittgenstein, *Ibid.*, 3.03).

Les linguistes et les logiciens — les premiers pour la langue naturelle, les seconds pour les langues formelles — expliquent la productivité de la langue par le fait que la signification est *compositionnelle* : la signification du tout est fonction de la signification des parties et de la manière dont celles-ci sont agencées.

La connaissance d'une langue exige, en outre, la capacité de comprendre et de produire un ensemble potentiellement infini de phrases. Cette deuxième caractéristique des langues est clairement décrite dans le texte suivant de *Aspects of The Theory of Syntax* de Chomsky :

> [...] la connaissance d'une langue met en jeu la capacité implicite de comprendre un nombre indéfini de phrases. De ce fait, une grammaire générative doit être un système de règles qui peuvent être itérées pour engendrer un nombre indéfiniment grand de structures (Chomsky, 1965, 1971, pp. 30–31).

Nous avons vu en 5.2.2 que de telles règles existent dans la syntaxe du calcul des prédicats. Ce sont les règles récursives de formation des termes et des formules. Nous verrons qu'elles doivent être doublées de règles récursives d'interprétation.

La première tâche de la sémantique du calcul des prédicats sera de dire à quel type d'objets renvoient les termes de base :

1. constantes individuelles;
2. symboles de prédicat;
3. symboles de fonction.

Ce qu'il y a de plus intuitif en sémantique extensionnelle, c'est certainement la relation de *dénotation* qui unit le nom propre d'un individu existant[1] à la chose dénotée. C'est cette relation que l'on *instaure* quand on baptise un enfant, une rue ou un bateau et que l'on *fait connaître* à quelqu'un quand on décline son identité.

La sémantique se donne pour tâche de montrer "comment le langage est ancré dans le monde (Smart, 1986, p. 98)". Pour accomplir cette tâche, on commence par se donner une *structure* $\mathcal{S}$.

**Définition 7.1 : structure d'interprétation d'un langage.** On appelle *structure d'interprétation* $\mathcal{S}$ du langage $\mathcal{L}$, ou plus simplement *interprétation* d'un langage $\mathcal{L}$, la paire $\langle \mathcal{D}, \mathcal{I} \rangle$, où $\mathcal{D}$ est un ensemble non vide, appelé *domaine* de $\mathcal{S}$, et $\mathcal{I}$ est une application[2], appelée *interprétation*, qui

1. envoie chaque constante individuelle $c$ sur un membre du domaine de $\mathcal{D}$, membre que l'on désigne par $c^{\mathcal{I}}$ (l'image de $c$ par l'application $\mathcal{I}$),
2. envoie chaque symbole de fonction (foncteur à $n$ places) $f$ sur une fonction qui prend ses arguments dans $\mathcal{D}^n$, et ses valeurs dans $\mathcal{D}$,
3. envoie chaque symbole de relation (prédicat à $n$ places) sur une relation $n$-aire, c'est-à-dire sur une classe de $n$-uples qui sont membres de $\mathcal{D}^n$.

Nous avons dit que l'immédiate intelligibilité des phrases nouvelles pouvait s'expliquer par le fait que la signification du tout est fonction de la signification des parties et de la manière dont celles-ci sont agencées. On croit parfois qu'il suffit de conférer une interprétation aux parties ultimes et que les règles syntaxiques de formation se chargeront de générer la signification du tout à partir de la signification des parties. C'est une erreur. Un francophone à qui on apprendrait dans un cours d'anglais ce qui suit :

1. "Peter" désigne l'individu Pierre;
2. "Mary" désigne l'individu Marie;
3. "(-) sees (-)" exprime la relation de voir;
4. en insérant des noms dans les places vides de (3), on génère une phrase anglaise;

---

[1] Si on prend "relation" au sens étroit où ce mot est utilisé en théorie des ensembles (une relation $n$-aire est une classe de $n$-uples), on ne peut dire que le nom propre "Pégase" dénote Pégase. En effet, seul le nom existe, pas la chose. Or la relation binaire de dénotation est une classe de couples ordonnés dont les deux membres existent. Voir J.J.C. Smart, 1989, "How to Turn the *Tractatus* Wittgenstein into (almost) Donald Davidson", in E. Lepore, *Truth and Interpretation*, Oxford, Blackwell § 186, p. 98.

[2] On appelle *application* une fonction *totale*, c'est-à-dire une correspondance qui à *chaque* élément $x$ du domaine $\alpha$ fait correspondre un élément du codomaine $\beta$. Soient les ensembles $\alpha = \{a, b, c\}$ du domaine et $\beta = \{1, 2, 3, 4\}$. La fonction $f$ telle que $a \rightarrow f(a) = 2$, $b \rightarrow f(b) = 1$ est une fonction partielle; mais si on y ajoute $c \rightarrow f(c) = 4$, la fonction $f'$ ainsi obtenue est une application.

serait incapable de dire si c'est Pierre qui voit Marie ou si c'est Marie qui voit Pierre. Il lui manque une règle *sémantique* de composition des significations qui *corresponde* à la règle *syntaxique* de composition des mots et qui précise la contribution *sémantique* de l'*ordre* des mots en français, c'est-à-dire la convention qui veut que l'on mette *devant* un verbe transitif l'agent et que l'on mette *derrière* le verbe le patient. En latin, cette différence est indiquée par les cas (*"Petrus videt Mariam"* a le même sens que *"Mariam videt Petrus"*). Comme le remarque Davidson,

> la connaissance des caractéristiques structurales [...] d'une phrase, plus la connaissance des significations des parties ultimes, ne s'additionnent pas pour donner comme résultat la connaissance de ce que la phrase signifie. [...] la syntaxe récursive à laquelle un dictionnaire a été ajouté n'est pas nécessairement une sémantique récursive (Davidson, 1967, 1969, p. 4).

Dans le calcul des prédicats, les expressions bien formées sont de deux sortes. Il y a quatre constructions grammaticales possibles : la *formation de termes* complexes, la *prédication*, l'*insertion de connecteurs* et la *quantification*. Aux règles de formation (R.F.) exposées en 5.2.2, nous devrons ajouter des règles sémantiques de composition des significations (R.S.).

## 7.3. Règles sémantiques pour les termes

Il existe trois sortes de termes :

1. les constantes individuelles telles que 0, 1, 2, Londres, Louis XIV, etc;
2. les variables individuelles telles que $x$, $y$, $z$, $x'$, $y'$, $z'$, etc;
3. les termes composés tels que le successeur de 0, le prédécesseur de Louis XIV, $2!$, $1+2$, $x+y$, $x+1$, $x+(y+z)$, etc.

Pour interpréter les constantes individuelles, on a recours à une *application* $\mathcal{I}$ appelée *interprétation*, qui fait correspondre à chaque terme singulier son référent dans l'univers du discours. Cette fonction $\mathcal{I}$ fait partie de la structure $\mathcal{S}$ au même titre que le domaine d'individus $\mathcal{D}$.

Pour interpréter les variables, on a recours à une *application* $\mathcal{A}$, distincte de $\mathcal{I}$, appelée *assignation*, qui fait correspondre à chaque variable un individu du domaine. Cette assignation ne fait pas partie de la structure $\mathcal{S}$ car une variable n'est pas un nom propre qui, en vertu de son sens, désignerait un individu particulier.

Pour interpréter les termes *hybrides* comme $x!$, on aura besoin à la fois de l'application $\mathcal{I}$ et de l'application $\mathcal{A}$. C'est pourquoi, la solution la plus simple est de dire que pour *tous* les termes on a besoin de $\mathcal{I}$ et de $\mathcal{A}$, mais que pour les constantes individuelles, on peut ignorer l'application $\mathcal{A}$ et pour les variables individuelles, l'application $\mathcal{I}$. Telle est la portée des deux règles R.S.1 et R.S.2 ci-dessous.

Les règles de formation des termes (R.F.1) comportent une règle récursive qui permet de former une infinité de termes (par exemple $S(0)$, $S(S(0))$, etc). Nous devrons donc associer à la règle récursive de formation une règle récursive d'interprétation. En effet, connaissant la signification [ici l'*extension*] de "successeur de (–)" et de "0", nous devons pouvoir générer — sans que l'on nous l'apprenne — la signification de "le successeur de 0", "le successeur du successeur de 0" et ainsi de suite.

La règle récursive de composition des significations qui permet d'obtenir ce résultat, quand on se borne aux constantes, peut s'écrire ainsi en français :

> L'interprétation du terme $[f(t_1, \ldots, t_n]$ = l'interprétation du terme $f$ appliquée aux interprétations des termes $t_1$ à $t_n$.

Par exemple,

> l'interprétation du terme $+(4!, 3^2)$ est l'interprétation du terme "+" appliquée à l'interprétation du terme "4!" et à l'interprétation du terme "$3^2$".
>
> Comme "4!" et "$3^2$" sont, à leur tour, des termes, on leur applique la règle d'abord :
>
> l'interprétation du terme "4!" est l'interprétation du terme "!" appliquée à l'interprétation du terme "4". On fait de même pour "$3^2$". Ensuite, on applique l'interprétation de "+" à celle de "4!" et de "$3^2$".

*Remarque.* Deux opérations sont à l'oeuvre ici :
l'opération syntaxique de *concaténation* $\star$, définie sur les symboles, et l'opération ensembliste d'*application fonctionnelle* $\circ$, définie sur les significations (extensionnelles) des symboles. La fonction *interprétation* $\mathcal{I}$ qui les relie satisfait la condition suivante :

$$(7.1) \qquad \|f \star (t_1, \ldots, t_n)\|^{\mathcal{I}} = \|f\|^{\mathcal{I}} \circ (\|t_1\|^{\mathcal{I}}, \ldots, \|t_n\|^{\mathcal{I}})$$

qui est une instance de :

$$(7.2) \qquad h(x \star y) = h(x) \circ h(y)$$

c'est-à-dire de la condition qu'une fonction $h$ doit satisfaire pour être un *homomorphisme*.

**Définition 7.2 : l'homomorphisme**[3]. Si un ensemble $E^1$ est muni d'une loi de composition interne $\star$, et si l'ensemble $E^2$ est muni d'une loi de composition interne $\circ$, une application $h$ de $E^1$ dans $E^2$ est un homomorphisme si, quels

[3] Soit $E^1$ l'ensemble des nombres complexes et $E^2$ l'ensemble des angles de direction. Soit l'opération $\cdot$ le produit défini sur $E^1$ et l'opération $+$ la somme définie sur $E^2$. Soit enfin $f$ une fonction qui prend ses arguments dans les nombres complexes représentés dans le plan de Gauss et ses valeurs dans les angles de direction; on a

$$f(v_1 \cdot v_2) = f(v_1) + f(v_2).$$

que soient $x$ et $y$, l'image par $h$ du composé $x \star y$ est le composé des images $h(x) \circ h(y)$.

Formulons à présent les règles sémantiques complètes pour les termes, en prenant en compte aussi les variables individuelles isolées ou incorporées à des symboles de fonction.

*Base.*

**RS1** Pour une constante d'individu $c$, $||c||_{\mathcal{I},\mathcal{A}} = ||c||_{\mathcal{I}}$. En d'autres termes, pour les constantes, l'assignation ne joue aucun rôle.

**RS2** Pour une variable $v$, $||v||_{\mathcal{I},\mathcal{A}} = ||v||_{\mathcal{A}}$. Pour les variables, l'interprétation ne joue auncun rôle.

*Récursion.*

**RS3** Pour un symbole de fonction $f$,

$$||f(t_1, \ldots, t_n)||_{\mathcal{I},\mathcal{A}} = ||f||_{\mathcal{I}}(||t_1||_{\mathcal{I},\mathcal{A}}, \ldots, ||t_n||_{\mathcal{I},\mathcal{A}}).$$

### 7.3.1. *La notion de satisfaction*

Nous avons vu que la signification des termes est compositionnelle. Nous devons à présent montrer qu'il en va de même pour les formules. Si les seules formules considérées étaient les propositions moléculaires formées au moyen des règles de formation 3.3.2 qui génèrent des propositions à partir de propositions atomiques en préfixant ou en insérant des connecteurs, la tâche serait aisée. En effet, l'extension d'une proposition étant une valeur de vérité et les connecteurs ($\neg$, $\wedge$, $\vee$, $\supset$) exprimant des fonctions de vérité, il est clair que l'extension du tout est fonction de l'extension des parties.

Pour interpréter les symboles propositionnels atomiques $P^1$, $P^2$, $P^3$,..., on assigne à ceux-ci une valeur de vérité.

La règle récursive de composition des valeurs de vérité, quand on se borne aux propositions moléculaires formées par l'insertion de $\wedge$, $\vee$, $\supset$, $\equiv$[4], peut s'écrire ainsi :

> L'interprétation de la formule $[g(X_1, X_2)]$ = l'interprétation du connecteur $g$ appliquée aux interprétations des formules $X_1$, $X_2$.

Par exemple,

> l'interprétation de $P^1 \vee (P^2 \supset P^3)$ est l'interprétation du connecteur $\vee$ appliquée à l'interprétation de $P^1$ (c'est-à-dire à une valeur de vérité) et à l'interprétation de $(P^2 \supset P^3)$.
> Comme $P^2 \supset P^3$ est à son tour une proposition composée (*moléculaire* selon la terminologie consacrée), on applique la règle à nouveau.

[4] Une définition du même type est possible pour le connecteur unaire $\neg$

Cette méthode échoue cependant si on tente de l'appliquer aux propositions générales de la forme $QxA$ où $Q$ est un quantificateur ($\exists$ ou $\forall$). En effet, sauf dans le cas où le quantificateur fonctionne *à vide*, la matrice $A$ d'une proposition générale n'est pas une proposition mais une fonction propositionnelle. Or une fonction propositionnelle n'a pas de valeur de vérité. On ne peut donc pas dire, par exemple, que :

> L'extension [valeur de vérité] de "$\exists x$ Pair($x$)" est fonction de l'extension [valeur de vérité] de "Pair($x$)".

Ce problème a été résolu par Tarski. Constatant qu'on ne peut aligner les fonctions propositionnelles sur les propositions et dire qu'elles *dénotent* comme ces dernières une valeur de vérité, il a fait l'inverse. Il a défini pour les fonctions propositionnelles la relation sémantique binaire de satisfaction (2.2.2) puis il a étendu celle-ci aux propositions (phrases fermées) et enfin il a défini la notion de vérité en termes de celle de satisfaction.

Ryle a justement critiqué les auteurs qui, assimilant toutes les relations sémantiques à celle de désignation, soutiennent que les phrases *nomment* une valeur de vérité et que les prédicats *nomment* une classe comme le nom propre "Fido" nomme le chien Fido. C'est ce qu'il a appelé la *"Fido"-Fido theory*. Tarski est à l'abri de ce reproche. Il reconnait que les fonctions propositionnelles telles que "$x$ est pair" ou "$x > y$" sont reliées à la réalité d'une autre manière que le nom "Fido" est relié au chien Fido. Une fonction propositionnelle à $n$ variables ne nomme pas une classe de $n$-uples ou une relation à $n$ termes mais elle est satisfaite ou non par une suite de $n$ termes du domaine du discours $\mathcal{D}$.

Tarski donne une *définition récursive* de la satisfaction dont Quine fait bien voir le mécanisme dans le passage suivant :

> Les énoncés peuvent être classés par ordre de complexité. Les prédications[5] ont la complexité 0; les négations et les quantifications existentielles d'énoncés de complexité $n$ sont de complexité $n + 1$; et les conjonctions ont une complexité $n + 1$ si l'un de leurs constituants a une complexité $n$ et leur autre constituant une complexité $n$ ou inférieure à $n$. (Quine 1970,1975 p. 65)

On commence par définir en quoi consiste la satisfaction pour un énoncé de complexité 0.

> Cela acquis, les conditions de la satisfaction pour la négation, la conjonction et la quantification existentielle nous disent en quoi consiste le fait qu'une suite satisfait un énoncé de complexité immédiatement supérieure une fois connu en quoi consiste le fait qu'une suite satisfait un énoncé d'une complexité donnée. (Quine, Ibid.)

[5] Quine considère un calcul des prédicats sans constantes individuelles. Les prédications seront donc toujours des phrases ouvertes ou fonctions propositionnelles de la forme $Fx$, $Rxy$, $Sxyz$, à l'exclusion de phrases fermées.

Dans notre définition récursive de l'interprétation $\mathcal{I}$ des formules, nous adopterons une terminologie différente qui est aujourd'hui plus courante. Au lieu de parler de fonctions propositionnelles *satisfaites* par des suites (ou plus exactement par des segments initiaux de suites), nous parlerons de fonctions propositionnelles *vraies relativement à une assignation* d'individus à leurs variables libres.

Selon cette nouvelle terminologie, nous dirons que

> La fonction propositionnelle "$x$ a conquis $y$" est vraie pour l'assignation de César à $x$ et de la Gaule à $y$ si et seulement si le couple ordonné ⟨César, la Gaule⟩ est membre de l'ensemble des couples ordonnés qui est l'extension du prédicat ("a conquis").

Nous arrivons alors à la définition récursive suivante [Fitting 1990, p. 105–106] de l'interprétation pour une formule.

**RS4** $||P(t_1, \ldots, t_n)||^{\mathcal{I},\mathcal{A}} = \mathbf{V}$ si et seulement si $\langle t_1^{\mathcal{I},\mathcal{A}}, \ldots, t_n^{\mathcal{I},\mathcal{A}} \rangle \in P^{\mathcal{I}}$,

**RS5** $||\neg X||^{\mathcal{I},\mathcal{A}} = \mathbf{V}$ si et seulement si $||X||^{\mathcal{I},\mathcal{A}} \neq \mathbf{V}$,

**RS6** $||X \wedge Y||^{\mathcal{I},\mathcal{A}} = \mathbf{V}$ si et seulement si $||X||^{\mathcal{I},\mathcal{A}} = \mathbf{V}$ et $||Y||^{\mathcal{I},\mathcal{A}} = \mathbf{V}$,

**RS7** $||\forall x \Phi||^{\mathcal{I},\mathcal{A}} = \mathbf{V}$ si et seulement si $\Phi^{\mathcal{I},\mathcal{B}} = \mathbf{V}$ pour toute assignation $\mathcal{B}$ dans la structure $\mathcal{S}$ qui est $x$-variante de $\mathcal{A}$, c'est à dire pour tout $\mathcal{B}$ identique à $\mathcal{A}$ sauf, *éventuellement*, pour ce qui est de l'individu assigné à $x$ par $\mathcal{B}$.

**RS8** $||\exists x \Phi||^{\mathcal{I},\mathcal{A}} = \mathbf{V}$ si et seulement si $\Phi^{\mathcal{I},\mathcal{B}} = \mathbf{V}$ pour au moins une assignation $\mathcal{B}$ dans $\mathcal{S}$ qui est une $x$-variante de $\mathcal{A}$.

La notion d'*$x$-variante* s'éclairera lors de l'application 7.3.2.

La fonction d'interprétation des formules ($\mathcal{I}$) est, comme celle des termes, un *homomorphisme*. Pour obtenir un *isomorphisme* (7.6.2), il aurait fallu exiger qu'il y eût une *bijection* entre les expressions bien formées et leur signification (ici leur extension). Or cette exigence est trop forte. Dans un langage bien fait, une même expression ne peut avoir deux extensions —sinon elle serait équivoque— mais deux expressions peuvent avoir la même extension, comme par exemple les deux prédicats "passagers du Mayflower" et "fondateurs de Plymouth, Massachusetts", qui ont pour extension la même classe.

La *vérité* peut être vue comme un cas limite de la satisfaction:

Une fonction propositionnelle peut être satisfaite par *toutes* les assignations, c'est la cas de

(7.3) $\quad x = x,$

par *au moins une* assignation, c'est la cas de

(7.4) $\quad x > y,$

ou par *aucune* assignation, c'est le cas de

(7.5) $\quad x > x.$

Au contraire, une proposition est satisfaite par *toutes les suites*, c'est le cas de

(7.6) $5 > 4$,

ou par aucune, c'est le cas de

(7.7) $4 > 5$.

La vérité est définie par Tarski comme étant le fait pour une proposition d'être satisfaite par toutes les suites, et, corrélativement, la fausseté est définie par le fait pour une proposition de n'être satisfaite par aucune suite.

En termes d'assignation plutôt que de suite, on aura:

**RS8** $\|A\|^{\mathcal{I}} = \mathbf{V}$ si $\|A\|^{\mathcal{I},\mathcal{A}} = \mathbf{V}$ pour toutes les assignations $\mathcal{A}$,

**RS9** $\|A\|^{\mathcal{I}} = \mathbf{F}$ si $\|A\|^{\mathcal{I},\mathcal{A}} = \mathbf{F}$ pour toutes les assignations $\mathcal{A}$.

La *validité* se définit comme la vérité dans toutes les interprétations, ce qui, formellement, se note $\models A$ si et seulement si $\|A\|^{\mathcal{I}} = \mathbf{V}$ pour tout $\mathcal{I}$. Par exemple,

(7.8) Si les cercles sont des figures géométriques, les cercles sont des figures géométriques

est valide parce qu'elle est vraie et reste vraie quelle que soit l'interprétation $\mathcal{I}$ que l'on donne à "cercle" et à "figure géométrique".

*Conclusion:* pour interpréter un langage comme le calcul des prédicats, on a besoin de trois choses

1. une structure $\mathcal{S} = \langle \mathcal{D}, \mathcal{I} \rangle$
2. une assignation $\mathcal{A}$
3. des règles de composition de sens **RS1** à **RS9**.

**Définition 7.3 : modèle.** On appelle modèle ($\mathcal{M}$) de $\Gamma$ une structure $\mathcal{S}$ qui rend vrai les propositions de l'ensemble $\Gamma$. Il ne faut pas confondre la relation "*rend vrai*" qui unit un modèle à des propositions avec la relation de *conséquence logique* qui unit des propositions à d'autres propositions. Pour marquer la différence nous utiliserons le symbole $\vDash$ dans le premier cas et le symbole $\models$ dans le deuxième.

### 7.3.2. *Application*

Pour interpréter la formule

(7.9) $\forall x \exists y Rxy \supset \exists x \forall y Rxy$

dans le modèle $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$ où

$$\mathcal{D} = \{\text{Albert}, \text{Bernard}, \text{Charles}\},$$
$$\mathcal{I} : R \mapsto \{\langle a, b\rangle, \langle a, c\rangle, \langle b, c\rangle, \langle b, a\rangle, \langle c, a\rangle, \langle c, b\rangle\},$$

on commence par analyser celle-ci au moyen d'un arbre syntaxique (figure 7.1), puis on assigne une interprétation aux noeuds terminaux, et on remonte ensuite dans l'arbre en appliquant les clauses de la définition récursive de la vérité.

Comme l'a dit joliment Quine :

le rapport entre une telle grammaire et la logique est que celle-ci explore les conditions de vérité des énoncés, à la lumière de la façon dont les énoncés sont grammaticalement construits. La logique pourchasse la vérité dans l'arbre de la grammaire (Quine, 1970, 1975, p. 57).

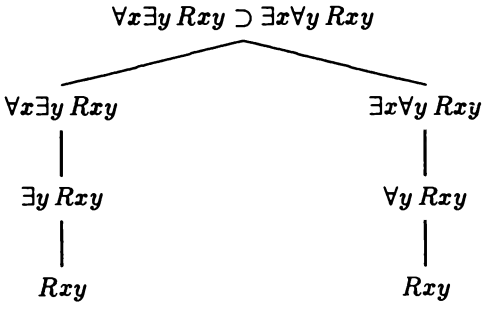



**Figure 7.1** *Arbre syntaxique de la formule 7.9*

$R$ pourrait être la relation exprimée par le prédicat de la langue naturelle "__ et ... se voient mutuellement". On suppose qu'il n'y a pas de miroir permettant à quelqu'un de se voir *lui-même* (cf. figure 7.2).

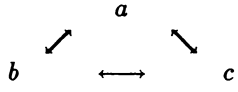



**Figure 7.2** *Graphe de la relation R*

Nous allons supposer que le domaine n'a que trois individus et nous limiter aux paires que l'on peut former avec ces derniers. Les 9 fonctions d'assignation ($\mathcal{A}$) possibles sont décrites à la figure 7.3.

Prenons l'une des 9 assignations, à savoir $\mathcal{A}_5$ c'est-à-dire $\langle b, b\rangle$ — le choix est sans importance —, et évaluons l'une des plus petites unités de sens de la formule, par exemple la phrase atomique ouverte $Rxy$.

L'assignation $\langle b, b\rangle$ ne satisfait pas $Rxy$. Satisfait-elle $\exists y Rxy$ ? Oui, s'il existe au moins une $y$-variante de $\langle b, b\rangle$ qui satisfait $Rxy$. Une $y$-variante de $\langle b, b\rangle$, c'est-à-dire de $\mathcal{A}_5$, c'est une suite qui diffère de $\mathcal{A}_5$ uniquement pour ce qui est de la valeur de $y$, et même pas nécessairement, car $\langle b, b\rangle$ est une $y$-variante d'elle-même.

Les $y$-variantes de $\langle b, b\rangle$ sont $\langle b, b\rangle$, $\langle b, a\rangle$ , et $\langle b, c\rangle$ . On voit que $\langle b, a\rangle$ et $\langle b, c\rangle$ vérifient $Rxy$ et donc que $\langle b, b\rangle$ vérifie $\exists Rxy$.

On cherche ensuite à savoir si l'assignation $\mathcal{A}_5$ satisfait $\forall x \exists y Rxy$. Pour cela, il faut que toutes les $x$-variantes d'au moins une $y$-variante de $Rxy$ satisfassent

| $g$ | $x$ | $y$ |
|---|---|---|
| 1 | $a$ | $a$ |
| 2 | $a$ | $b$ |
| 3 | $a$ | $c$ |
| 4 | $b$ | $a$ |
| 5 | $b$ | $b$ |
| 6 | $b$ | $c$ |
| 7 | $c$ | $a$ |
| 8 | $c$ | $b$ |
| 9 | $c$ | $c$ |

**Figure 7.3** *Liste des assigantions*

$Rxy$. On a vu que notamment l'$y$-variante $\langle b, a\rangle$ de la paire $\langle b, b\rangle$ satisfaisait $Rxy$. On cherche toutes les $x$-variantes de $\langle b, a\rangle$ : ce sont respectivement la suite $\langle b, a\rangle$ elle-même, ainsi que $\langle c, a\rangle$ et $\langle a, a\rangle$. Les deux premières assignations satisfont $Rxy$, mais pas la troisième. En effet, il n'y a pas de boucle sur le point $a$ dans le graphe. Mais ce qui est requis, ce n'est pas que toutes les $x$-variantes de $\langle b, a\rangle$ satisfassent $Rxy$, mais seulement que, pour toutes les $x$-variantes, il existe au moins une $y$-variante de $\langle b, a\rangle$ — pas nécessairement la même chaque fois — qui satisfasse $Rxy$. Or il existe une $y$-variante de $\langle a, a\rangle$ qui satisfait $Rxy$, à savoir $\langle a, b\rangle$. En résumé, on a

$\langle b, a\rangle$, une $x$-variante d'elle-même,
$\langle c, a\rangle$, une $x$-variante de $\langle b, a\rangle$, et
$\langle a, b\rangle$, une $x$-variante d'une $y$-variante de $\langle b, a\rangle$,

qui, toutes trois, satisfont $Rxy$. L'assignation $\mathcal{A}_5$ satisfait $\forall x \exists y Rxy$.

On refait le même travail pour $\exists x \forall y Rxy$. Ici, de nouveau, $\langle b, b\rangle$ échoue à satisfaire $Rxy$. De plus, il n'est pas vrai que toutes les $y$-variantes de $\langle b, b\rangle$ satisfassent $Rxy$. En particulier, l'$y$-variante (dégénérée) de $\langle b, b\rangle$, à savoir $\langle b, b\rangle$ elle-même, échoue à satisfaire $Rxy$. On examine ensuite les $x$-variantes de $\langle b, b\rangle$ pour voir s'il n'y en a pas au moins une qui satisfasse $\forall y Rxy$. On constate qu'aucune des autres $x$-variantes de $\langle b, b\rangle$ ne remplit cette condition. En effet, ni $Ray$ n'est vrai de tout $y$ : on n'a pas $\langle a, a\rangle$, ni $Rcy$ n'est vrai de tout $y$ : on n'a pas $\langle c, c\rangle$. Donc $\exists x \forall y Rxy$ n'est pas satisfaite dans le modèle utilisé ici. En effet, il s'agit d'une conditionnelle dont l'antécédent est satisfait mais dont le conséquent ne l'est pas. pour exprimer cela, formellement, on écrit :

(7.10) $\mathcal{M} \not\models \forall x \exists y Rxy \supset \exists x \forall y Rxy$.

## 7.4. Correction du calcul des prédicats

Nous nous bornerons à prouver que tous les théorèmes du système axiomatique présenté en 5.3 sont des formules valides. La preuve se fait par induction sur la longueur des démonstrations (Hamilton).

*Clause de base* : Si $\Phi$ a été obtenue par une démonstration à une ligne, $\Phi$ est un axiome. Or les axiomes sont valides (**A.1** à **A.3** sont des tautologies du calcul des propositions. La validité de **A.4** et **A.5** sera établie plus bas).

*Clause inductive* : Si $\Phi$ a été obtenue par une démonstration de $n > 1$ lignes, on suppose par hypothèse inductive que tous les théorèmes dont la démonstration a moins de $n$ lignes sont valides. Si $\Phi$ apparaît dans une démonstration, soit $\Phi$ est un axiome — alors $\Phi$ est valide (cas de base) — soit $\Phi$ a été obtenue à partir de formules antérieures dans la démonstration, par une application du *modus ponens* ou de la *règle de génération universelle*. Les prémisses de ces deux règles sont valides par hypothèse inductive (puisqu'elles figurent dans une démonstration ayant $k$ lignes où $k < n$. Pour prouver la validité de $\Phi$, il suffit de s'assurer que les deux règles transportent la validité. (Pour la première règle, voir 3.4.6.2 Pour la deuxième, voir ci-dessous.)

### 7.4.1. *Validité des axiomes propres au calcul des prédicats*

**A.4** $\forall\nu(\Psi \supset \Phi) \supset (\Psi \supset \forall\nu\Phi\nu)$ où $\nu$ est une variable individuelle ne figurant pas à titre d'occurrence libre dans $\Psi$.

*Démonstration* (d'après Hamilton 1978, p73)

Soient une interprétation $\mathcal{I}$ et une assignation $\mathcal{A}$ telles que $\| \forall\nu(\Psi \supset \Phi) \|^{\mathcal{I},\mathcal{A}} = \mathbf{V}$; alors en vertu de la définition de la satisfaction, toutes les assignations $\mathcal{B}$ qui ne diffèrent de $\mathcal{A}$ que pour ce qui est de $\nu$, si toutefois elles le font (c'est-à-dire pour toutes les $\nu$-variantes), satisfont $\Psi \supset \Phi$. Or, en vertu de la définition de la satisfaction, cela signifie que ces assignations satisfont $\Phi$ ou ne satisfont pas $\Psi$. Si elles ne satisfont pas $\Psi$, $\| \Psi \supset \forall\nu\Phi\nu \|^{\mathcal{I},\mathcal{A}} = \mathbf{V}$ en tant que conditionnel dont l'antécédent est faux (relativement à $\mathcal{I}$ et $\mathcal{A}$). Si pour toutes les $\nu$-variantes $\| \Phi\nu \|^{\mathcal{I},\mathcal{A}} = \mathbf{V}$, alors $\| \forall\nu\Phi\nu \|^{\mathcal{I},\mathcal{A}} = \mathbf{V}$ en vertu de la définition de la satisfaction et $\| \Psi \supset \forall\nu\Phi\nu \|^{\mathcal{I},\mathcal{A}} = \mathbf{V}$ en tant que conditionnel dont le conséquent est vrai (relativement à $\mathcal{I}$ et $\mathcal{A}$).

**A.5** $\forall\nu\Phi \supset \Phi_{\tau/\nu}$

La démonstration par une induction est assez longue. Nous renvoyons à Rivenc ou à Hamilton.

### 7.4.2. *Règle dérivée*

Si la variable $\nu$ ne figure à titre d'occurrence libre dans aucune des formules membres de l'ensemble $\Gamma$, alors

$$\text{Si } \Gamma \vdash \Phi \text{ alors } \Gamma \vdash \forall \nu \Phi$$

*Démonstration* (d'après Stoll 1961, 1963, p. 391–392)

Si $\Gamma \vdash \Phi$, il y a une suite finie de formules membres de $\Gamma$, à savoir $\Psi_1, \ldots, \Psi_n$ telles que

$$\Psi_1, \ldots, \Psi_n \vdash \Phi$$

ce qui peut s'écrire

$$\Psi_1 \wedge \ldots \wedge \Psi_n \vdash \Phi$$

Comme $\nu$ n'apparaît libre dans aucune formule $\Psi_i$,on peut appliquer le théorème de la déduction et obtenir

$$\vdash \Psi_1 \wedge \ldots \wedge \Psi_n \supset \Phi$$

Par généralisation universelle, on dérive

$$\vdash \forall \nu((\Psi_1 \wedge \ldots \wedge \Psi_n) \supset \Phi)$$

Par application de l'axiome **A.4**, on en tire

$$\vdash (\Psi_1 \wedge \ldots \wedge \Psi_n) \supset \forall \nu \Phi$$

Puis, en utilisant la version suivante du *modus ponens* (cf. 3.4.11),

(7.11) $$\frac{\vdash X \supset Y}{X \vdash Y}$$

on arrive à

$$\Psi_1 \wedge \ldots \wedge \Psi_n \vdash \forall \nu \Phi$$

et après remplacement des $\wedge$ par "," à

$$\Gamma \vdash \forall \nu \Phi$$

## 7.5. La vérité chez Aristote et chez Tarski

On dit de croyances et de propositions qu'elles sont vraies ou fausses. On le dit aussi de dents ou de revolvers, mais alors c'est dans un sens dérivé. On a en vue le contraste entre le *naturel* et l'*artificiel*, ou l'*authentique* et l'*imité*. Comme les croyances peuvent s'exprimer en propositions, on peut se borner à prendre les propositions (au sens d'expressions linguistiques; cf. Gochet, 1972) comme porteur du vrai et du faux.

Aristote définit la notion du vrai et du faux de la manière suivante :

> Dire de l'Etre qu'il n'est pas, ou du Non-Etre qu'il est, c'est le faux; dire de l'Etre qu'il est, et du Non-Etre qu'il n'est pas, c'est le vrai... (Aristote, Métaphysique, Γ.7.25).

Cette définition est le *locus-classicus* de la conception de la vérité connue sous le nom de la *vérité-correspondance* (cf. S. Haack et P. Engel).

La théorie de la vérité-correspondance est souvent reformulée ainsi : "la vérité d'une proposition réside dans sa conformité aux faits". La définition de la vérité en termes de satisfaction, proposée par Tarski, est souvent perçue comme une version *nouvelle* de la théorie de la vérité-correspondance. Davidson y voit en outre une version *améliorée* de cette dernière. Le progrès, selon lui, réside en ceci qu'elle permet de mettre en correspondance avec la réalité non seulement les propositions, porteuses attitrées du vrai et du faux, mais aussi les fonctions propositionnelles, privées de cette prérogative. Cela tient au fait que

> [...] la satisfaction des propositions [*closed sentences*] est exprimée en termes de la satisfaction des fonctions propositionnelles [*open sentences*] et des propositions, alors que, traditionnellement, seules les propositions correspondent à des faits (Davidson, 1969, 1984, p. 48).

Dans le *Vocabulaire technique et critique de la philosophie* de Lalande (édit. 1951), la vérité est présentée comme rebelle à toute définition :

> [...] les deux notions corrélatives et inséparables du vrai et du faux sont au nombre des idées essentielles et primitives sans lesquelles rien ne peut être pensé et qui, par conséquent, ne sont pas susceptibles d'une définition proprement dite (Lalande, 1951, p. 1224).

C'était aussi l'opinion de Descartes qui, dans une lettre à Mersenne du 16 octobre 1639, écrit :

> on peut bien... dire que le mot de vérité... dénote la conformité de la pensée avec l'objet... Mais on n'en peut donner aucune définition de la logique qui aide à connaître sa nature.

Nous partageons avec Davidson l'opinion que la définition de Tarski reprend et améliore celle d'Aristote. Doit-on, pour autant, attribuer à Tarski le mérite d'avoir accompli une tâche que Descartes jugeait irréalisable ? Popper l'a peut-être pensé quand il a prononcé cette affirmation dithyrambique lors de la célébration du soixante-dixième anniversaire de Tarski :

> Je souhaite rappeler ma joie intense et mon soulagement quand j'appris en 1936 que les choses suivantes résultaient de la théorie de la vérité de Tarski : que ce concept était définissable en termes logiques que personne n'avait envisagés auparavant [...] (Popper 1972, cité par Delahaye, 1990).

La réponse doit être nuancée. S'il s'agissait d'isoler l'*extension* (ou la dénotation) du prédicat "vrai" appliqué à une langue formalisée, il est indiscutable que Tarski y est parvenu. S'il s'agissait de donner l'*intension* (ou

le sens) du prédicat "vrai", il n'y est nullement parvenu. Putnam a établi ce point de manière irréfutable :

> La propriété à laquelle Tarski donne le nom de "vrai-dans-L" est une propriété que la phrase "La neige est blanche" possède dans tous les mondes où la neige est blanche, y compris dans ceux où cela signifie que la neige est verte. Du point de vue des applications formelles de la théorie de Tarski en logique mathématique, cela n'a pas d'importance puisque ce que veulent les logiciens d'une définition de la vérité est qu'elle capture l'extension (la dénotation) de "vrai" quand on l'applique à L, et non qu'elle capture le sens de la notion intuitive de vérité (restreinte à L). Mais la préoccupation de la philosophie est précisément de découvrir ce qu'est la notion intuitive de vérité. *En tant qu'analyse de la vérité, la théorie de Tarski échoue autant que cela est possible pour une analyse d'échouer.* [...] Vraiment, le nom que Tarski donna à sa théorie, "La Conception Sémantique de la Vérité", est gravement trompeur ! Il aurait été plus exact de l'appeler "La Conception Non Sémantique de la Vérité" (Putnam, 1985, pp. 61–79, cité par Delahaye, 1990).

Pour apprécier équitablement l'apport de Tarski, nous devons comparer le résultat à l'objectif visé. Or l'objectif[6] de Tarski n'était pas de définir le *sens* du mot "vrai" (par opposition à son extension), mais de montrer que l'on pouvait utiliser le mot "vrai" sans risquer d'introduire de paradoxe. Le but premier visé par Tarski était, comme le rappelle opportunément Etchmendy,

> de fournir une définition éliminative de la vérité, de manière à "absorber" le concept dans une métathéorie ne contenant aucun concept sémantique ... (Etchemendy, 1988, pp. 52–53).

Il est clair qu'on a besoin du terme "vrai" pour énoncer certains problèmes scientifiques; par exemple, pour poser la question résolue négativement par Gödel, "Toutes les propositions *vraies* de l'arithmétique intuitive sont-elles dérivables comme théorèmes dans un système axiomatique formalisé ?" Or ce concept, qui semble essentiel,

> [...] menace d'introduire des inconsistances dans toute discipline qui l'emploie (Etchemendy, 1988, p. 53, cité par Delahaye, 1990).

Tarski a réussi à conjurer ce danger. Le paradoxe du menteur, que certains usages du mot "vrai" peuvent générer, est écarté des langues formelles qui se plient aux exigences de Tarski. Il faut cependant, ici aussi, relativiser la portée de son résultat, pour important qu'il soit. D'autres manières de donner une

[6] Dans "La conception sémantique de la vérité" (Tarski, 1944, 1974, p. 280), Tarski écrit : "En particulier, nous voulons que les *termes sémantiques* (relatifs au langage-objet) *ne soient introduits dans le métalangage qu'au moyen de définitions*". Les développements ultérieurs de la sémantique formelle et notamment l'emploi de la notion sémantique de *vérité relative à une assignation* (cf. 7.2 et 7.3) ont fait perdre de vue l'objectif premier de Tarski.

définition extensionnelle de la vérité, et de "bloquer" le paradoxe du menteur, ont été proposées depuis, et celles-ci constituent un progrès par rapport à la solution de Tarski, en cela qu'elles sont moins mutilantes (elles bannissent moins de propositions comme étant mal formées).

## 7.6. Vérité et signification

Une sémantique compositionnelle doit rendre possible un "calcul du sens", selon l'heureuse expression de M. Galmiche (1990, à paraître). Or quand nous avons calculé, de proche en proche, l'interprétation de la formule 7.9 de la section 7.3.2, nous avons obtenu non pas une signification, mais une valeur de vérité. A première vue, ce n'est pas ce qu'on attendrait d'une sémantique.

La *compétence sémantique* consiste, en effet, à pouvoir assigner une *signification* aux propositions — par exemple à la phrase "Il y a de la vie sur Mars" — tandis que la *compétence encyclopédique* consiste à pouvoir assigner à celles-ci une *valeur de vérité*.

Certes, la sémantique que nous venons d'examiner est purement extensionnelle; elle ne prend pas en compte l'intension, mais se borne à assigner aux symboles une extension. Or si dans le cas des termes et des fonctions propositionnelles, l'extension est, respectivement, l'objet dénoté et une classe de n-uples, dans le cas des propositions, l'extension est précisément la valeur de vérité. On pourrait néanmoins nous objecter d'avoir subrepticement remplacé la thèse intuitivement satisfaisante 1.6.5 exprimée par l'équation

(7.12) signification de la proposition $p$ = conditions de vérité de $p$,

par l'équation inacceptable

(7.13) signification de la proposition $p$ = valeur de vérité de $p$.

La réponse à cette critique est la immédiate : ce que la définition récursive de la vérité dans un modèle nous permet de faire, c'est de calculer la valeur de vérité d'une formule dans un modèle donné *pour autant que l'on sache quels individus dans le domaine du modèle satisfont les fonctions propositionnelles*. Elle ne s'arroge nullement le pouvoir exorbitant de trancher a priori des questions touchant la réalité. Il faut, pour s'en convaincre, garder présente à l'esprit la condition formulée dans l'énoncé en italique. Dans la vie quotidienne, cette condition n'est pas remplie. Nous connaissons plus souvent l'intension (le sens) des termes généraux (noms communs, prédicats) que leur extension. Le théatre de boulevard exploite abondamment le fait que nous connaissons le sens du mot "père", mais pas son extension.

La définition de Tarski ne nous permet pas de calculer la valeur de vérité de la proposition "Quelques philosophes sont ingénieurs" sans plus. Elle nous permet seulement de calculer la valeur de vérité de cette proposition dans un modèle dans lequel l'extension du prédicat "philosophe" et celle du prédicat "ingénieur" ont été fixées au départ. Remarquons que celui qui sait que la classe

constituant l'extension de "philosophe" contient comme membres, entre autres, Thalès et Wittgenstein, et qui sait aussi que la classe constituant l'extension de "ingénieur" contient aussi ces membres, celui-là montrerait qu'il ne comprend pas le *sens* de la phrase précitée s'il disait ne pas en connaître la valeur de vérité. La définition récursive de la vérité dans un modèle est donc en parfait accord avec notre intuition.

Nous venons de montrer que la définition récursive de la vérité dans un modèle n'efface pas la distinction entre *valeur de vérité* et *signification* d'une proposition. Mais, si elle ne s'occupe que de l'extension, comment peut-elle rendre compte positivement de la différence entre deux propositions qui ont même valeur de vérité, mais qui ont des sens différents, comme c'est le cas pour "la terre attire la lune" et "la lune attire la terre" ? La réponse est claire : on dira que ces deux propositions *disent* quelque chose de différent l'une de l'autre en cela qu'elles sont vraies dans des modèles différents, ce qui n'exclut pas qu'elles soient vraies *aussi* dans le même modèle (celui qui représente notre monde). Nous avons donc ici une manière *purement extensionnelle* de caractériser la *signification* d'une proposition : la signification d'une proposition, c'est la classe des modèles dans lesquels elle est vraie. Hintikka a bien mis ce point en lumière :

> Selon la conception défendue de manière cohérente par la théorie des modèles, ce qu'une proposition (*sentence*) $S$ dit, elle le dit en spécifiant une classe de modèles $\mathcal{M}(S)$ [...] De ce point de vue, il est clair qu'une proposition [...] fournit d'autant plus d'information qu'elle restreint davantage la classe des modèles dans lesquels elle est vraie. Elle n'est pas d'autant plus informative qu'elle s'applique à plus de modèles (qu'elle est vraie en ceux-ci), mais au contraire, elle est d'autant plus informative que la classe des modèles dans lesquels elle ne s'applique pas (dans lesquels elle est fausse) est grande (Hintikka, 1988, p. 4)[7].

### 7.6.1. *Axiomes, modèles et contenu d'information*

La réponse de Hintikka appelle un commentaire. On pourrait se demander comment le contenu d'information d'une théorie $T$ peut consister dans la classe des modèles dans lesquels elle est vraie. La question qui vient tout de suite à l'esprit est la suivante : Si les propositions de la théorie $T$ ne sont pas dotées d'une *signification préalable*, qu'est ce qui restreint la classe des modèles dans lesquels les propositions de $T$ sont vraies ? La réponse est que la théorie $T$ contient des *axiomes propres* qui ne sont pas vrais dans tous les modèles,

[7] Hintikka parle ici des modèles $\mathcal{M}$ qui rendent vrais les axiomes d'une *théorie* et non des structures $\mathcal{S}$ qui se bornent à conférer une signification aux expressions bien formées d'un langage.

contrairement aux axiomes de la logique. Ces axiomes extralogiques *isolent* une classe de modèles : la classe des modèles qui les rendent vrais. Nous allons le montrer au moyen d'un exemple.

Notons ici que lorsqu'on fait varier l'interprétation, c'est-à-dire que l'on change le domaine $\mathcal{D}$ et l'interprétation $\mathcal{I}$, on laisse cependant intacts les constantes logiques de la théorie. Celles-ci conservent leur signification standard (à moins que la théorie $T$ ne soit précisément une logique, mais nous laisserons ce cas de côté pour l'instant).

Pour comprendre l'affirmation, déconcertante à première vue, selon laquelle la signification des propositions provient de la valeur de vérité que celles-ci possèdent dans un modèle, alors qu'on s'attend à ce que le sens *précède* la valeur de vérité, il est utile de réfléchir sur un cas concret.

Considérons le système axiomatique suivant :

(7.14)
1. $F(a)$,
2. $\forall x(F(x) \supset F(x'))$,
3. $\forall x \forall y((x' = y') \supset (x = y))$,
4. $\forall x \neg(x' = a)$,
5. $(\Phi(a) \wedge \forall x(\Phi(x) \supset \Phi(x'))) \supset \forall x \Phi x$.

Les constantes logiques $\{\supset, \neg, \forall, =\}$ reçoivent leur interprétation usuelle. Les axiomes comportent trois constantes non logiques (le terme singulier $a$, le prédicat unaire $F$, et enfin le symbole de fonction unaire $'$). Le symbole $\Phi$ est un marque-place de formule ouverte de complexité quelconque. Alors que les axiomes 1 à 4 sont des propositions non interprétées, l'axiome 5 est un *schéma d'axiome* qui engendre un axiome dès que l'on substitue une formule ouverte au marque-place $\Phi$.

Un des modèles visés est le suivant : $\mathcal{D} = \mathbf{N}$ et $\mathcal{I}$ envoie $a$ sur le nombre zéro, $F$ sur l'ensemble des naturels, et $'$ sur la fonction *successeur immédiat de* (c'est-à-dire $x' = x + 1$).

Ce modèle vérifie les axiomes, mais le modèle suivant, donné par B. Russell dans *Introduction to Mathematical Philosophy* (1919, p. 7), le ferait aussi : $\mathcal{D} = \{1, 1/2, 1/4, 1/8, 1/16 \ldots\}$ et $\mathcal{I}$ envoie $a$ sur 1, $F$ sur $\mathcal{D}$ et $'$ sur la fonction *moitié de*.

En revanche, le modèle suivant est exclu : $\mathcal{D}$ est l'ensemble des entiers ($\mathcal{Z}$), $a$ est 0 et $'$ signifie *successeur*. En effet, 0 est précédé des membres de $\{-1, -2, \ldots\}$.

Un système d'axiomes agit comme un tamis. Il élimine certains modèles et par cette capacité d'exclusion, il confère *indirectement* une signification à ses constantes non logiques (ici $a$, $F$, $'$) et un contenu d'information aux axiomes. D'Euclide à Frege et Russell, les systèmes d'axiomes ont joué un rôle uniquement épistémologique. *Ils servaient les besoins de la preuve*. Ils avaient pour rôle de fournir un point de départ aux démonstrations. Avec l'avènement de la *théorie des modèles*, ils jouent un rôle nouveau : ils permettent de décrire des structures et de définir des notions primitives d'une manière contextuelle. La

logique du premier ordre, dans laquelle l'univers $\mathcal{D}$ est un ensemble d'individus, suffit à représenter toutes les *preuves* dont on fait usage en mathématiques. Mais comme on confère un rôle nouveau aux systèmes d'axiomes — celui de *décrire des structures* plutôt que de fournir des méthodes de preuve — la question se pose de savoir si la logique du premier ordre suffit encore, compte tenu de ce rôle nouveau, ou s'il faut recourir à une logique du deuxième ordre dans laquelle $\Phi$ cesse d'être un marque-place et est promu au rang de variable de prédicat susceptible d'être liée par un quantificateur ($\forall\Phi$). Nous reviendrons plus loin sur les enjeux d'une réponse à cette question (voir fin de la section 9.2, le début de la section 9.3.5, théorème 9.13).

### 7.6.2. *Catégoricité et vérité*

La catégoricité est la propriété d'une théorie dont tous les modèles sont isomorphes. La notion technique d'isomorphisme vise à traduire l'idée intuitive de *structure commune* dont on se sert dans la vie quotidienne.

> Par exemple, le patron en papier d'une robe de coton peut être dit isomorphe au découpé de la robe pour ce qui est de la dimension et de la forme. Il n'en diffère que par le matériau (Partee, 1978, p. 11).

**Définition 7.4 : isomorphisme.** Soient le modèle $\mathcal{M}_1$ constitué d'un domaine $\mathcal{D}_1$, muni d'une relation $\mathcal{R}_1$ et d'une opération $\mathcal{O}_1$ définie sur $\mathcal{D}_1$, et le modèle $\mathcal{M}_2$ constitué d'un domaine $\mathcal{D}_2$ et d'une opération $\mathcal{O}_2$. On dit que $\mathcal{M}_1$ est isomorphe à $\mathcal{M}_2$ si les conditions suivantes sont remplies :

- $f : \mathcal{D}_1 \mapsto \mathcal{D}_2$ où $f$ est une bijection
- $\langle x, y\rangle \in \mathcal{R}_1$ si et seulement si $\langle f(x), f(y)\rangle \in \mathcal{R}_2$
- $f(x\mathcal{O}_1 y) = f(x)\mathcal{O}_2 f(y)$

Un système axiomatique formalisé décrit bien les modèles, c'est-à-dire les interprétations dans lesquelles les axiomes sont vrais, mais on ne peut pas dire qu'il décrit ce qu'il y a de commun à tous les modèles qui le vérifient, car il peut n'y avoir aucun *contenu* commun :

> [...] étant donné n'importe quelle interprétation $\mathcal{I}_i$, il y a d'autres interprétations isomorphes à $\mathcal{I}_i$, mais qui n'ont pas de contenu en commun avec $\mathcal{I}_i$. L'existence de telles "images" isomorphes implique, évidemment, l'impossibilité de caractériser de manière unique une interprétation au moyen d'un ensemble de propositions (*sentences*) dans une langue formelle. En conséquence, la meilleure des caractérisations possibles d'une interprétation serait une caractérisation à un isomorphisme près (Corcoran, 1980, p. 190, cité par Shapiro 1985, p. 718).

Ceci permet de préciser la différence entre les vérités de mathématique et les vérités de physique. Comme l'écrit G. Hirsch,

> Une théorie physique vise la description, schématisée et simplifiée d'ailleurs, de situations complexes et, de ce fait, uniques en leur genre. Si la théorie n'est pas adéquate pour la situation considérée, il est à peu près vain d'espérer rencontrer autre part des situations qui seront décrites par ladite théorie. La mathématique, au contraire, étudie des situations moins complexes et plus schématisées [...]; ces structures ont donc de plus grandes chances de se retrouver en diverses occasions [...] (Hirsch, 1975, p. 9).

G. Hirsch illustre le contraste par la comparaison suivante :

> Si en dessinant le plan d'une petite ville du Middle West, où seulement deux ou trois rues se croisent à angle droit, le cartographe se trompe et permute, par exemple, la banque et l'école, il se trouvera certainement d'autres petites villes pour lesquelles le plan pourra quand même servir — cas des structures mathématiques et même de la géométrie euclidienne; si le plan était celui de nos villes européennes et s'il présente une erreur, il ne se trouvera aucune ville à laquelle s'appliquerait le plan ainsi modifié — cas de la physique (Hirsch, *Ibid.*, p. 8).

Dans son *Introduction to logic* (1941, trad. fr. 1971), Tarski esquisse une théorie en miniature qui illustre bien le trait des langues formelles mis en évidence par Corcoran. Cette théorie comporte les deux axiomes

(7.15) $\forall x R(x,x)$,

(7.16) $\forall x \forall y \forall z((Rxz \wedge Ryz) \supset Rxy)$,

d'où il est facile de déduire les théorèmes

(7.17) $\forall y \forall z(Ryz \supset Rzy)$,

(7.18) $\forall x \forall y \forall z((Rxy \wedge Ryz) \supset Rxz)$.

Parmi les modèles qui vérifient les axiomes et donc aussi les théorèmes, on a les deux structures suivantes:

- $\mathcal{M}_1 = \langle \mathcal{D}_1, \mathcal{I}_1 \rangle$, où $\mathcal{D}_1$ est l'ensemble des *segments*[8], et où $\mathcal{I}_1$ envoie le prédicat binaire $R(-,-)$ sur la relation de *congruence*[9] entre segments.
- $\mathcal{M}_2 = \langle \mathcal{D}_2, \mathcal{I}_2 \rangle$, où $\mathcal{D}_2$ est l'ensemble des nombres rationnels, et $\mathcal{I}_2$ envoie le prédicat binaire sur la relation d'*équivalence* entre nombres définie ainsi : les nombres $x$ et $y$ sont équivalents — ce qui se note $x \equiv y$ — si et seulement si $x - y = z$ où $z$ est un entier. On peut donc affirmer que $1\frac{1}{4} \equiv 5\frac{1}{4}$, mais pas que $3 \equiv 2\frac{1}{3}$.

Il est clair qu'entre les modèles $\mathcal{M}_1$ et $\mathcal{M}_2$, il n'y a pas de *contenu* commun.

Non seulement les théories du premier ordre ne peuvent caractériser une interprétation de manière unique, mais elles ne peuvent même pas en général

[8] On appelle *segment* la partie d'une droite comprise entre deux points de celle-ci.

[9] La congruence entre segments, c'est la superposabilité.

caractériser à un isomorphisme près un modèle, sauf dans le cas particulier d'un modèle fini. Cependant, même pour une théorie qui n'est pas catégorique, il arrive que l'on puisse établir qu'un sous-ensemble intéressant de ses modèles existe tel que

> n'importe quel modèle de la théorie est isomorphe à quelque membre de ce sous-ensemble (Suppes, 1957, p. 263).

La possibilité de l'application des mathématiques[10] au monde physique repose sur l'existence de cette forme affaiblie de catégoricité.

> Etant donné une théorie axiomatisée de la mesure de quelque quantité empirique, la tâche qui incombe au mathématicien est de démontrer un théorème de représentation[11] pour les modèles de la théorie qui établit, pour dire les choses sommairement, que n'importe quel modèle est isomorphe à quelque modèle numérique de la théorie (Suppes, 1957, p. 266).

Dans "A Set of Independent Axioms for Extensive Quantities" (1951, 1969), P.Suppes présente un système d'axiomes pour la théorie empirique de la pesée à l'aide d'une balance à fléau. Le modèle visé (*"intended model"*) comporte un domaine d'objets physiques individuels que l'on peut placer sur un des plateaux de la balance. Ces objets sont ordonnés par la relation *plus léger ou égal en poids à* et on peut définir sur ce domaine l'opération de *réunir sur un même plateau de la balance*.

Pour pouvoir appliquer les nombres à ces objets, on définit un modèle intermédiaire où le domaine est formé de nombres réels positifs, ordonnés par la relation $\leq$. L'opération de réunir sur un même plateau de la balance correspond ici à l'addition définie sur les nombres en question.

Pour relier le modèle physique au modèle mathématique, on définit une relation d'équivalence $C$ telle que $xCy = xQy \ \& \ yQx$ et où "$xQy$" signifie "$x$ est plus léger ou égal en poids à $y$". Cette relation permettra de diviser l'ensemble des objets pesables en classes d'équivalence groupant les objets de même poids. Ensuite on prouve le *théorème de représentation* suivant dans lequel la notion d'isomorphisme joue un rôle important :

> Si $\mathcal{M} = \langle K, Q, \star \rangle$ est un système de grandeurs extensives, alors la partition obtenue en prenant l'ensemble quotient (en divisant $K$ par la relation

[10] Pour dire les choses simplement, si on veut formuler non seulement des *affirmations qualitatives* comme "cet objet $a$ est plus lourd que cet objet $b$", mais aussi des *affirmations quantitatives* comme "cet objet $a$ pèse 50 grammes de plus que cet objet $b$", c'est parce qu'il est possible d'établir que "la structure d'un ensemble de phénomènes sous certaines opérations et relations empiriques est *la même* que la structure de quelque ensemble de nombres fermé sous certaines opérations et relations arithmétiques" (Suppes *Ibid.* Cf. aussi Suppes 1951, 1969).

[11] Le théorème de représentation le plus connu est le *théorème de Stone*, qui affirme que toute algèbre de Boole est isomorphe à un sous-ensemble de parties $\mathcal{P}(U)$ muni des opérations de réunion, d'intersection, de complémentation et de 2 éléments distingués, la classe vide et la classe universelle.

*C*) est isomorphe à un demi-groupe additif de nombres réels, fermé sur la soustraction des plus petits à partir des plus grands.

## 7.7. La conséquence logique

### 7.7.1. *La conséquence logique comme relation*

On se souvient de la définition qu'Aristote a donnée du syllogisme (section 1.3.4) : "le syllogisme est un discours dans lequel, certaines choses étant posées, quelque chose d'autre que ces données en résulte nécessairement". Que faut-il entendre par "résulter nécessairement" ?

Aujourd'hui, comme nous savons que la logique permet de reproduire sous forme de preuves formalisées les raisonnements exacts des mathématiciens, nous sommes tentés de répondre que la proposition $C$ *résulte nécessairement* des propositions $P_1, \ldots, P_n$ si et seulement s'il existe une *dérivation formelle* de $C$ à partir des prémisses $P_1$ à $P_n$, ce qui se note $P_1, \ldots, P_n \vdash C$. La notion de *dérivation* est une notion de syntaxe logique qui a été définie rigoureusement par Carnap dans *Logische Syntax der Sprache* (1934, 1964, p. 28) :

> Par *dérivation* avec prémisses $\sigma_1, \sigma_2, \ldots, \sigma_m$ (dont le nombre est toujours fini, et peut être 0), nous entendons une série de propositions (*sentences*) de n'importe quelle longueur finie, telle que chaque proposition de la série est soit l'une des prémisses, soit une proposition-définition, soit directement dérivable de l'une ou de plusieurs [...] des propositions qui la précèdent dans la série (Carnap, 1934, 1964, p. 28).

Une proposition $\sigma_3$ est alors définie par Carnap comme directement dérivable de la proposition $\sigma_1$ ou des propositions $\sigma_1$ et $\sigma_2$ si elle peut être obtenue à partir de $\sigma_1$ ou de $\sigma_1$ et $\sigma_2$ par des règles d'inférence immédiate comme la substitution ou le *modus ponens*, règles que Carnap appelle "règles de transformation" pour souligner le fait que l'on ne tient compte que de la structure externe des propositions quand on les applique.

Malheureusement, il existe des théories mathématiques — notamment l'arithmétique et toute théorie qui la contient — pour lesquelles la notion de *dérivation* est trop étroite pour recouvrir celle de conséquence logique. Gödel a, en effet, démontré que "dans chaque système dans lequel les mathématiques peuvent être formulées, des propositions peuvent y être construites qui sont valides au sens des mathématiques classiques, mais qui ne sont pas démontrables dans le système (Carnap, *Ibid.*, p. 100)".

Ceci a conduit Carnap à introduire un concept de *conséquence* distinct[12] du

[12]La première différence entre conséquence logique et dérivation, c'est que l'ensemble des prémisses dans le cas de la relation de conséquence peut être un ensemble infini. La deuxième différence, c'est que dans une chaîne de conséquences, même les maillons individuels ne sont

concept de ***dérivabilité***, mais néanmoins syntaxique. Carnap voit même dans celui-ci "le point de départ de toute syntaxe" et il ajoute :

> Si pour un langage quelconque le terme "conséquence" a été établi, alors tout ce qui peut être dit concernant les connexions logiques dans ce langage est par là-même déterminé (Carnap, *Ibid.*, p. 168).

Carnap définit la conséquence logique pour un certain langage assez riche pour exprimer les mathématiques classiques. Il le fait à l'aide de la notion syntaxique de *contradiction* :

> la proposition $X$ ***suit logiquement*** des propositions de la classe $K$ si et seulement si la classe composée de toutes les propositions de $K$ et de la négation de $X$ est contradictoire (Carnap, *Ibid.*, p. 117, formulation de Tarski 1936, 1974, p. 147).

Cette définition syntaxique de la notion de conséquence logique fait intervenir la négation. Elle n'est pas assez générale. On ne pourrait l'appliquer à un langage qui serait dépourvu de négation, comme le fragment implicationnel du calcul des propositions qui contient uniquement le connecteur $\supset$.

Dans le *Tractatus logico-philosophicus* (1922), Wittgenstein propose une définition de la conséquence logique que l'on peut rétrospectivement qualifier de sémantique. Appelons "bases de vérité" les assignations de valeurs de vérité aux atomes propositionnels qui rendent vraie une proposition construite avec des connecteurs vérifonctionnels :

> [...] la vérité d'une proposition $p$ découle de la vérité d'une autre proposition $q$ si toutes les bases de vérité de cette dernière sont contenues dans les bases de vérité de la première (Wittgenstein, 1922, 5.12).

Par exemple, la proposition $A \wedge B$ a, parmi ses conséquences logiques, $A \vee B$ parce que toutes les distributions de valeurs de vérité qui rendent la première vraie $\{\mathcal{I}_1\}$ ***sont contenues*** dans celles qui rendent la deuxième vraie $\{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3\}$ (cf. figure 7.4).

La définition de Wittgenstein convient parfaitement pour le calcul des propositions. Ne peut-on l'étendre aux propositions du calcul des prédicats qui, outre les atomes propositionnels et les connecteurs, contiennent des prédicats,

---

pas "définis". Carnap précise le contraste entre dérivation et conséquence logique en ces termes :

> Une dérivation est une suite finie de propositions, une série de conséquences est une série finie de classes qui ne sont pas nécessairement finies. Dans le cas des dérivations, chaque pas individuel (c'est-à-dire la relation "directement dérivable") est effectif (*definite*), mais pas la relation "dérivable", qui est définie (*defined*) par la chaîne entière des dérivations. Dans le cas d'une série de conséquences, le pas immédiat (la relation de "conséquence directe") est déjà non effectif (*indefinite*) et *a fortiori* la relation de "conséquence" (Carnap, *Ibid.*, p. 39).

| | $A$ | $B$ | $A \wedge B$ | $A \vee B$ |
|---|---|---|---|---|
| $\mathcal{I}_1$ | **V** | **V** | **V** | **V** |
| $\mathcal{I}_2$ | **V** | **F** | **F** | **V** |
| $\mathcal{I}_3$ | **F** | **V** | **F** | **V** |
| $\mathcal{I}_4$ | **F** | **F** | **F** | **F** |

**Figure 7.4** *Tableau matriciel*

des foncteurs, des variables et des quantificateurs ? Si l'on se borne aux propositions et que l'on écarte les fonctions propositionnelles, l'extension est immédiate. Une proposition $B$ est une conséquence logique d'un ensemble éventuellement infini de propositions $\Gamma$ si et seulement si tous les modèles qui rendent vrais les membres de $\Gamma$ rendent vraie la proposition $B$.

Nous représentons la relation sémantique de conséquence logique à l'aide du signe $\models$, qu'il faut distinguer du *turnstile*, c'est-à-dire du signe $\vdash$, lequel représente la relation syntaxique de dérivabilité. $A \models B$ signifie "$B$ est conséquence logique de $A$", tandis que $A \vdash B$ signifie "$B$ est dérivable de $A$ par les règles du système". Ces deux relations relient des propositions, c'est-à-dire des formules fermées, ou des ensembles de propositions. Munis de cette notation, nous pouvons formuler la définition sémantique de la *conséquence logique* pour les formules fermées.

**Définition 7.5 : la conséquence logique pour les formules fermées.**

$\Gamma \models B$ si et seulement si pour tout $\mathcal{S}$, si $\|\Gamma\|_{\mathcal{S}} = \mathbf{V}$, alors $\|B\|_{\mathcal{S}} = \mathbf{V}$.

Par exemple, on a

(7.19) $\quad \{p, p \supset q\} \models q$

En effet toute interprétation $\mathcal{I}$ de $p$ et de $q$ dans $\mathcal{D} = \{\mathbf{V}, \mathbf{F}\}$ qui rend vrais les membres de $\Gamma$, ici $p$ et $q$ rendent vrai $B$, c'est-à-dire $q$.

De même, on a

(7.20) $\quad p \models q \vee \bar{q}$

En effet, il n'y a pas d'interprétation qui rende la proposition $p$ vraie et la proposition $q \vee \bar{q}$ fausse, cette dernière étant une tautologie, c'est-à-dire une proposition vraie pour toutes les interprétations.

Les concepts sémantiques dont on a besoin pour définir la conséquence logique entre formules en général (c'est-à-dire entre formules ouvertes ou fermées) sont plus compliqués. Dans une interprétation $\mathcal{I}$ donnée, une formule fermée est vraie ou fausse. Ce n'est plus le cas pour les fonctions propositionnelles. Certes, on peut appliquer, même aux fonctions propositionnelles, le prédicat "vrai dans l'interprétation $\mathcal{I}$". Ainsi, on dira que la fonction

propositionnelle $\Phi x$ est vraie si sa clôture universelle $\forall x \Phi x$ est vraie et que la même fonction propositionnelle $\Phi x$ est fausse si sa clôture existentielle $\exists x \Phi x$ est fausse. Mais, contrairement à ce qui était le cas pour les propositions, il reste une troisième possibilité, celle dans laquelle une fonction propositionnelle n'est pas vraie ou fausse dans une interprétation donnée, mais où elle est dans l'interprétation en question, vraie pour certaines assignations de valeur à ses variables libres (ici à $x$) et fausse pour d'autres.

Considérons le fragment de langage $\mathcal{L}$ constitué par les expressions

$$\{x, \Phi(-), \Phi x, \forall x \Phi x, \supset, \Phi x \supset \forall x \Phi x\}$$

et les quatre structures $\mathcal{S}_1 = \langle \mathcal{D}, \mathcal{I}_1 \rangle$, $\mathcal{S}_2 = \langle \mathcal{D}, \mathcal{I}_2 \rangle$, $\mathcal{S}_3 = \langle \mathcal{D}, \mathcal{I}_3 \rangle$ et $\mathcal{S}_4 = \langle \mathcal{D}, \mathcal{I}_4 \rangle$, où $\mathcal{D} = \{1, 2\}$.

| | $\Phi$ | $x$ | $\Phi x$ | $\forall x \Phi x$ | $\Phi x \supset \forall x \Phi x$ |
|---|---|---|---|---|---|
| $\mathcal{I}_1$ | $\{1,2\}$ | 1 | V | V | V |
| | $\{1,2\}$ | 2 | V | V | V |
| $\mathcal{I}_2$ | $\{1\}$ | 1 | V | F | F |
| | $\{1\}$ | 2 | F | F | V |
| $\mathcal{I}_3$ | $\{2\}$ | 1 | F | F | V |
| | $\{2\}$ | 2 | V | F | F |
| $\mathcal{I}_4$ | $\{\ \}$ | 1 | F | F | V |
| | $\{\ \}$ | 2 | F | F | V |

**Figure 7.5** *Relation de conséquence et calcul des prédicats*

Toutes les interprétations $\mathcal{I}_1$ qui rendent vraie la formule ouverte $\Phi x$ rendent vraie la formule fermée $\forall x \Phi x$ et toutes les interprétations $\mathcal{I}_4$ qui rendent fausse la formule fermée $\forall x \Phi x$ rendent fausse la formule ouverte $\Phi x$. Mais, contrairement aux formules fermées, une formule ouverte comme $\Phi x$ peut, pour une même interprétation $\mathcal{I}$ de la constante $\Phi$, être vraie pour une assignation de la variable $x$ et fausse pour une autre. Cette situation est illustrée dans $\mathcal{I}_2$ et $\mathcal{I}_3$ de la figure 7.5.

Largeault dit la même chose, mais en termes de *suites* :

> Puisque le vrai [...] est la satisfaction par toutes les suites, il est exact qu'il n'y a pas d'interprétation qui rend "$\forall x P x$" faux et "$P x$" vrai; néanmoins "$P x$" n'implique pas [n'a pas pour conséquence logique] "$\forall x P x$", car il n'est pas exact que toute suite qui satisfait "$P x$" satisfait aussi "$\forall x P x$". (Largeault 1980, p. 86)

Une définition de la notion de conséquence logique qui convient aux formules quelconques (aussi bien ouvertes que fermées) devra être plus restrictive que la définition 7.5. Bien qu'il n'existe pas d'*intreprétation* qui rende vrai $\Phi x$ et faux $\forall x \Phi x$ (vu la définition de la vérité d'une fonction propositionnelle donnée plus haut), il existe des *assignations* de valeurs à $x$ qui, combinée à

une interprétation de $\Phi$ rendent vrai $\Phi x$ et faux $\forall x \Phi x$. En d'autres termes, le conditionnel $\Phi x \supset \forall x \Phi x$ n'est pas valide.

**Définition 7.6 : conséquence logique pour les formules quelconques.** La proposition $B$ est une conséquence logique (au sens restreint) de l'ensemble de propositions ou de fonctions propositionnelles $\Gamma$ si et seulement si, dans toute interprétation $\mathcal{I}$, toute assignation $\mathcal{A}$ qui satisfait simultanément les membres de $\Gamma$ satisfait aussi $B$.

*Remarque.* On voit ici l'importance de la distinction entre *interprétation* et *assignation*.

### 7.7.2. *Trois sens du mot "complétude"*

Tant que l'on restait enfermé dans la perspective universaliste de l'idéographie de Frege et des *Principia mathematica* de Whitehead et Russell, on ne pouvait envisager la complétude qu'au sens "expérimental", c'est-à-dire se demander si toutes les preuves couramment utilisées par les mathématiciens pouvaient être reproduites dans la langue formelle.

Avec la dissociation entre langage et métalangage, il devient possible de faire la théorie de notions comme celle de dérivation et de conséquence. Si l'on définit la conséquence en termes syntaxiques comme on l'a fait pour la dérivabilité, ce que fait Carnap, on peut certes voir la différence entre les deux notions pour les *systèmes incomplets*, dans lesquels il y a des propositions qui sont des conséquences des axiomes sans en être dérivables, mais la différence s'estompe pour les systèmes complets et corrects où toutes les conséquences sont dérivables et réciproquement. La notion de *conséquence* et celle de *dérivabilité* risquent de se confondre.

Au contraire, une fois que l'on dispose de la distinction entre syntaxe et sémantique établie par Tarski, il devient aisé d'expliquer ce qu'est la complétude. Le prédicat métalinguistique à deux places "__ est dérivable de..." et le prédicat métalinguistique à deux places "__ est conséquence logique de..." ne sont jamais *synonymes*, ils n'ont jamais la même *intension*. L'un se définit en termes de syntaxe, de symboles et de manipulation de symboles. L'autre se définit en partie en termes de sémantique, c'est-à-dire de relations unissant les symboles à un *au-delà des symboles*, à une réalité[13] généralement non-linguistique. S'il est vrai que les notions de conséquence et de dérivation ne sont jamais *synonymes*, elles peuvent cependant être *coextensives*.

[13] Même dans l'interprétation de Herbrand, il faut distinguer les constantes "$a$", "$f(a)$", "$f(f(a))$", ...dans leur rôle de *symboles* et ces mêmes constantes dans leur rôle de *référents* (comme il faut distinguer "mot" comme signifiant de "mot" comme signifié).

**Définition 7.7 : la complétude forte.** Nous dirons qu'un système est *fortement complet* au sens sémantique si et seulement si à toute relation de conséquence entre une proposition $B$ et un ensemble de propositions $\Gamma$ correspond une relation de dérivabilité :

(7.21) Si $\Gamma \models B$ alors $\Gamma \vdash B$.

**Définition 7.8 : la complétude faible.** De la complétude forte, on doit distinguer la *complétude faible* (au sens sémantique) qui est un cas dégénéré de la première forme de complétude, le cas où l'ensemble $\Gamma$ est vide :

(7.22) Si $\models B$ alors $\vdash B$,

ce qui peut se lire : "un système est faiblement complet si toutes les propositions logiquement valides sont des théorèmes".

Le calcul des prédicats possède ces deux formes de complétude. La relation syntaxique désignée par $\vdash$ et la relation sémantique désignée par $\models$ y sont en parfaite correspondance, y compris pour ce qui est de l'introduction du conditionnel. Les restrictions imposées au théorème de la déduction et à la définition de $\models$ pour les formules quelconques interdisent respectivement, par exemple,

(7.23) $$\frac{Fx \vdash \forall x Fx}{\vdash Fx \supset \forall x Fx}$$

et

(7.24) $$\frac{Fx \models \forall x Fx}{\models Fx \supset \forall x Fx}$$

mais autorisent, par exemple,

(7.25) $$\frac{Fx \vdash (Fx \vee Gx)}{\vdash Fx \supset (Fx \vee Gx)}$$

et

(7.26) $$\frac{Fx \models (Fx \vee Gx)}{\models Fx \supset (Fx \vee Gx)}$$

**Définition 7.9 : la complétude syntaxique.** Il existe encore un autre sens du mot complétude, c'est celui de *complétude syntaxique*. Un système axiomatique est syntaxiquement complet si et seulement si toute formule bien formée du système $S$ ou sa négation est démontrable, ce qui se note

(7.27) $\vdash_S A$ ou $\vdash_S \neg A$, pour toute formule fermée $A$.

La catégoricité entraîne la complétude syntaxique et la complétude syntaxique entraîne à son tour la complétude sémantique, mais la réciproque n'est pas vraie (pour la démonstration de cette chaîne d'implications, voir Martin 1965, pp. 89-90). Rappelons aussi que la complétude syntaxique (mais non la complétude sémantique) entraîne la décidabilité (voir section 2.4.5).

Le calcul des prédicats du premier ordre possède les deux premières formes de complétude — cela sera démontré plus loin — mais pas la troisième. La proposition suivante, par exemple, n'est pas valide :

(7.28) $\exists x Px \supset \forall x Px$.

Elle n'est vraie que dans les modèles qui ont un domaine à un membre. Sa négation n'est pas valide non plus : elle n'est vraie que dans les modèles qui ont un domaine à $n$ membres, où $n > 1$.

### 7.7.3. *Calcul des prédicats et théorie du premier ordre*

Le calcul des prédicats du premier ordre et les théories du premier ordre ont en commun le fait de n'utiliser comme variables que des variables d'individus : on s'interdit de lier par un quantificateur des variables de prédicats ou de fonctions. Les axiomes logiques sont également communs au calcul des prédicats du premier ordre et aux théories du premier ordre. Ce qui les différencie, c'est d'abord le langage. Alors que dans le calcul des prédicats du premier ordre les symboles de prédicats et de fonctions ne reçoivent pas d'interpétation déterminée, dans les théories du premier ordre certains prédicats ou symboles de fonctions sont isolés des autres et reçoivent une signification par le truchement d'axiomes propres à la théorie, c'est-à-dire d'axiomes extralogiques.

Les axiomes du calcul des prédicats sont valides, c'est-à-dire vrais pour tous les modèles. Au contraire, les axiomes d'une théorie du premier ordre ne sont vrais que pour le ou les modèle(s) qu'elle est supposée décrire. Pour souligner le contraste, on pourrait dire que les axiomes du calcul des prédicats sont *valides*, tandis que les axiomes propres d'une théorie du premier ordre sont *vrais*. Il ne faut cependant pas insister exagérément sur cette opposition. On doit en effet se rappeler que les théories coulées en systèmes axiomatiques formalisés ne sont pas vraies ou fausses absolument, mais ne sont vraies que pour certains modèles.

La distinction entre complétude sémantique et complétude syntaxique ne suffit pas à dissiper l'équivoque que l'usage de ce terme comporte. Il conviendrait encore de distinguer une *version sémantique* et une *version syntaxique* de la complétude dite syntaxique.

La version syntaxique de la complétude syntaxique est très clairement exprimée dans le passage suivant :

> Une théorie formalisée est appelée *complète* si pour chaque formule fermée dans le langage de cette théorie qui représente une proposition, c'est-à-dire pour chaque formule dans laquelle n'apparaît aucune variable libre, il y a une preuve soit de cette proposition, soit de sa négation (et donc il n'y a pas de proposition indécidable). (Rasiowa, 1973, p. 284)

La version sémantique de la complétude syntaxique pourrait se formuler ainsi : "une théorie est complète si toute formule vraie dans un modèle de la

théorie est vraie dans tous". La théorie de l'ordre linéaire dense et non borné en est un exemple. Ces deux versions ne sont pas synomymes, mais elles sont coextensives.

Toute théorie du premier ordre est sémantiquement complète au sens où le calcul des prédicats est sémantiquement complet. En effet, toute théorie du premier ordre est une particularisation du calcul des prédicats qui hérite des propriétés de ce calcul. En revanche, il existe des théories du premier ordre qui ne sont pas syntaxiquement complètes, ni dans la version syntaxique ni dans la version sémantique de cette dernière notion. La théorie des groupes est de cette sorte. Elle contient des énoncés tels que

$$\forall x \forall y (x.y = y.x)$$

qui sont vrais dans certains modèles et faux dans d'autres, tout comme la formule du calcul des prédicats

$$\exists x Fx \supset \forall x Fx$$

était réalisable (*satisfiable* en anglais) dans certaines interprétations et falsifiable dans d'autres.

Si l'on ne prend pas soin de faire les distinctions présentées ici, on ne comprendra pas qu'il soit possible d'affirmer simultanément, sans se contredire :

1. les théories du premier ordre sont sémantiquement complètes,
2. le système axiomatique formalisé de l'arithmétique dû à Peano est une théorie du premier ordre,
3. le système axiomatique de Peano est incomplet.

Le système axiomatique de l'arithmétique de Peano est complet en ce sens que toutes les propositions vraies dans tous les modèles de la théorie axiomatisée sont démontrables dans cette théorie. Il est incomplet en ce sens que que toutes les propositions vraies dans l'interprétation standard (dans **N**) du langage de l'arithmétique ne sont pas démontrables dans la théorie axiomatique de Peano[14]. L'ensemble des vérités de l'arithmétique non axiomatisée n'est pas récursivement énumérable. Il contient parmi ses parties l'ensemble des théorèmes de l'arithmétique formalisée, ensemble qui est récursivement énumérable.

Pour nous résumer, les mots "complétude" et "incomplétude" qui figurent respectivement dans le théorème de Gödel de 1930 (le calcul des prédicats du premier ordre — et les théories du premier ordre qui en sont des extensions, telle que l'arithmétique formalisée — est complet) et dans le théorème de Gödel de 1931 (l'arithmétique formalisée et toutes les théories axiomatiques qui l'incluent sont incomplètes), ne sont pas employés dans le même sens.

[14] Cette incomplétude est irrémédiable. On ne peut la supprimer en choisissant d'autres axiomes (aussi longtemps, bien entendu, qu'on impose à l'ensemble des axiomes l'obligation d'être récursif (2.3.3)).

### 7.7.4. *Validité et conséquence logique*

On peut définir la *validité* à partir de la *conséquence* logique et dire qu'une formule valide est une formule qui est la conséquence logique de la classe vide des prémisses. On peut également définir la validité — et c'est peut-être plus intuitif — à partir de la *vérité* : une formule est *valide* si et seulement si elle est vraie dans *tous* les modèles. On donne le nom de *vérité de logique* aux propositions qui sont des instances de formules valides dans ce sens.

Deux autres concepts sémantiques doivent être introduits : la *consistance sémantique* (en anglais, *satisfiability*) et l'*inconsistance sémantique* (en anglais, *unsatisfiability*). Une proposition est *sémantiquement consistante* si et seulement si elle est vraie dans *quelque* modèle, et *sémantiquement inconsistante* si elle ne l'est dans *aucun*.

L'exemple suivant illustre la différence entre proposition vraie dans quelques modèles, mais pas tous, et proposition valide, c'est-à-dire vraie dans tous les modèles. La proposition

(7.29) $\forall x \exists y Ryx$

est vraie dans le modèle $\mathcal{M}_1 = \langle \mathcal{D}_1, \mathcal{I}_1 \rangle$, où $\mathcal{D}_1 = \mathbf{N}$, c'est-à-dire l'ensemble des nombres naturels, et $\mathcal{I}_1$ assigne au prédicat binaire $R(--)$ la relation *successeur de* définie sur le produit cartésien $\mathcal{D}_1 \times \mathcal{D}_1$. Elle est vraie aussi dans le modèle $\mathcal{M}_2 = \langle \mathcal{D}_2, \mathcal{I}_2 \rangle$, où $\mathcal{U}_2 = \mathbf{R}$ et $\mathcal{I}_2$ envoie $R(-,-)$ sur la relation *est plus grand que* (définie sur le produit cartésien $\mathcal{D}_2 \times \mathcal{D}_2$). En revanche, elle est fausse pour $\mathcal{M}_3 = \langle \mathcal{D}_3, \mathcal{I}_3 \rangle$, où $\mathcal{D}_3 = \{1, 2, 3\}$ et $\mathcal{I}_3$ est la relation *successeur de* (ou *plus grand que*) définie sur le produit cartésien $\mathcal{D}_3 \times \mathcal{D}_3$.

Au contraire, la formule

(7.30) $\exists x \forall y Rxy \supset \forall y \exists x Rxy$

est vraie dans tous les modèles $\mathcal{M}$, c'est-à-dire pour tous les choix de $\mathcal{D}$ et de $\mathcal{I}$. Elle est donc valide et toute proposition qui en est une instance est logiquement vraie.

Nous avons vu que l'on peut définir la *validité* comme cas-limite de la *conséquence logique*. Peut-on faire l'inverse, et définir la conséquence logique à l'aide du concept de validité ? Ne pourrait-on dire que $B$ est une conséquence logique de l'ensemble $\Gamma$ de formules fermées si et seulement si le *conditionnel* $(\wedge\Gamma) \supset B$ formé par la conjonction[15] des formules contenues dans $\Gamma$ prise comme antécédent et $B$ pris comme conséquent est un conditionnel valide ? La réponse est affirmative *seulement si* l'ensemble $\Gamma$ est *fini*. Dans ce dernier cas, on peut en effet dire que

(7.31) $\Gamma \models B$ si et seulement si $\models (\wedge\Gamma) \supset B$,

mais on ne peut plus le faire quand $\Gamma$ est infini (Etchemendy, 1988, p. 76). *On*

[15] On fait ici le même abus d'écriture que dans la définition de conséquence logique donnée en 7.7.1, et on convient que $\wedge\{a, b, c, \ldots\}$ est mis pour $a \wedge b \wedge c \ldots$.

*doit donc conclure que la notion de conséquence logique est plus puissante que celle de vérité logique.*

Frege, Russell et Quine lui-même défendent une vue de la logique selon laquelle la logique serait constituée d'un corps de vérités — les vérités de la logique — comme l'arithmétique est constituée d'un corps de vérités, les vérités de l'arithmétique.

Si on définit la logique à l'aide de la notion de *conséquence logique* plutôt qu'à l'aide de la notion de *vérité logique* — ce choix étant justifié par le fait que la notion de conséquence est plus puissante — alors la notion de vérité de logique apparaît comme "un cas dégénéré de la relation de conséquence logique" :

> Les vérités de la logique sont les vérités qui sont les conséquences de n'importe quelles prémisses ou d'aucune. En particulier, elles ne dépendent pas d'un ensemble propre d'axiomes... (Etchemendy, *Ibid.*, p. 74),

contrairement aux vérités de mathématique par exemple.

La définition de la logique fondée sur la notion de conséquence logique a également pour corollaire

> ...la vue selon laquelle la logique n'est pas simplement une théorie de plus, sur le même pied que l'arithmétique ou la géométrie, qui traite d'un corps distinctif de vérités, mais qu'elle est plutôt une métadiscipline qui traite des théories déductives en général (Etchemendy, *Ibid.*, p. 75).

Si on met à l'avant-plan la notion de *conséquence logique* plutôt que celle de *vérité logique*, on mettra à l'avant-plan le correspondant syntaxique de la notion de conséquence logique, c'est-à-dire la relation de dérivabilité, et les règles d'inférence, c'est-à-dire de déduction, dans lesquelles la relation de dérivabilité se matérialise.

C'est cette conception de la logique, comme *métadiscipline*, que nous adopterons. Les axiomes de la logique ne seront pas traités comme des axiomes à mettre sur le même pied que les axiomes de l'arithmétique, mais seront traités comme des règles d'inférence déguisées ou dégénérées. Pour prendre un exemple concret, nous traiterons l'axiome

(7.32) $\vdash p \supset (p \vee q)$

comme dérivé, par rapport à la règle d'inférence

(7.33) $p \vdash (p \vee q)$,

qui peut se lire "De $p$ déduisez $p \vee q$".

Puisque nous suivons Etchemendy et considérons la relation sémantique de *conséquence logique* et la relation syntaxique de *dérivabilité* comme les premiers sujets de la logique, il importe d'étudier à présent ces notions pour elles-mêmes et de les caractériser de la manière la plus générale possible.

### 7.7.5. *Axiomatiser la relation de conséquence logique*

Est une relation de conséquence logique ($\models$) toute relation entre formules qui vérifie les axiomes suivants :

1. *Axiome de réflexivité.*

(7.34) $\qquad B \models B.$

2. *Axiomes de monotonie.* On peut augmenter les prémisses par *conjonction* ou augmenter les conclusions par *disjonction*[16], ce qui revient, au fond, à les diminuer :

(7.35) $\qquad$ Si $A_1, \ldots, A_n \models B$ alors $A_1, \ldots, A_n, C \models B,$

(7.36) $\qquad$ Si $A_1, \ldots, A_n \models B$ alors $A_1, \ldots, A_n \models B, C,$

3. *Axiome de transitivité.*

(7.37) $\qquad$ Si $A \models B$ et $B \models C$ alors $A \models C.$

Jusqu'ici, la relation binaire de conséquence logique est définie sur un *ensemble* de formules $\mathcal{E}$ non-interprétées, à savoir sur le produit cartésien de $\mathcal{E}$ par $\mathcal{E}$.

Elle n'est pas encore une relation sémantique. Pour qu'elle le devienne, il faut associer à l'ensemble de formules $\mathcal{E}$ une interprétation *exhaustive* (toutes les formules membres de $\mathcal{E}$ ont une interprétation) et *cohérente* (aucune formule ne reçoit deux interprétations différentes). En logique bivalente, une interprétation est une fonction qui prend ses arguments dans $\mathcal{E}$ et ses valeurs dans l'ensemble des valeurs de vérité $\{\mathrm{V}, \mathrm{F}\}$:

(7.38) $\quad \mathcal{I} : \mathcal{E} \mapsto \{\mathrm{V}, \mathrm{F}\}.$

Si le cardinal de $\mathcal{E}$, c'est-à-dire le nombre d'éléments de $\mathcal{E}$ est $n$, il y a $2^n$ interprétations possibles. On peut donc écrire :

> La conséquence logique est incluse dans le produit cartésien de $2^{\mathcal{E}}$ par lui-même.

Jusqu'ici, les formules membres de $\mathcal{E}$ sont des formules atomiques. Quand on introduira des connecteurs, l'ensemble $\mathcal{E}$ s'enrichira de propositions moléculaires obtenues à partir de formules atomiques combinées avec des connecteurs, puis à partir de formules moléculaires combinées avec des connecteurs, et ainsi de suite... Appelons $\mathcal{E}^*$ le plus petit ensemble de formules contenant $\mathcal{E}$ et fermé sous les connecteurs $\wedge, \vee, \supset$ (Scott, 1973, p. 258). L'ensemble $\mathcal{E}^*$ sera construit inductivement à partir de $\mathcal{E}$ (à l'aide de la définition récursive de formule, qui contient des clauses telles que "si $A \in \mathcal{E}^*$ et $B \in \mathcal{E}^*$, alors $A \wedge B \in \mathcal{E}^*$").

[16] La virgule a un sens différent selon qu'elle est à gauche ou à droite du symbole $\models$ (cf. 4.1.1)

La fonction d'interprétation $v$ sera étendue en $v^*$ par une définition récursive qui contient des clauses telles que

(7.39) $\mathcal{I}^*(A \wedge B) = \mathbf{V}$ si et seulement si $\mathcal{I}^*(A) = \mathbf{V}$ et $\mathcal{I}^*(B) = \mathbf{V}$.

On doit ensuite démontrer que la relation de conséquence $\models^*$ étendue maintenant à l'ensemble $\mathcal{E}^*$ continue à satisfaire les trois axiomes ainsi que les règles d'inférence associées aux connecteurs telles que, par exemple,

(7.40) $(A \wedge B) \vdash A$ et $(A \wedge B) \vdash B$.

Ces considérations abstraites sur la relation de *conséquence logique* nous permettent de transcender les particularités de la logique classique :

> L'analyse en termes de valeurs de vérité n'est pas tributaire de l'adoption de connecteurs semblables aux connecteurs booléens standard. La question de savoir quels sont les connecteurs autorisés et quelles sont les propriétés de fermeture qu'ils peuvent posséder doit toujours être posée [...] On peut imaginer des discussions qui exigeraient que l'on fasse des distinctions sur le "degré" et la "profondeur" de l'itération des connecteurs (Scott, *Ibid.*, pp. 261–262).

On peut, par exemple, définir un connecteur "si...alors..." qui ne transporterait pas la vérité intégralement, mais qui autoriserait une dérive limitée par rapport à la vérité. Supposons qu'il y ait des degrés de vérité entre 0 (vérité maximum) et $n$ (erreur) et que le conditionnel transporte la vérité au prix d'un accroissement d'erreur $i$. Soit, par exemple, $A \rightarrow B$, un conditionnel qui a le degré de vérité $i$. Supposons que l'antécédent $A$ ait le degré de vérité $j$. Le conséquent $B$ devra seulement avoir le degré de vérité de $i + j$ (Scott, *Ibid.*, pp. 271–272).

L'exemple suivant montre l'utilité de ce conditionnel ($\rightarrow$) qui diffère du conditionnel philonien ($\supset$). Soit $\approx$ la relation d'indiscernabilité de deux points. En raison des seuils de perception, la distance entre les deux points $a$ et $b$, notée $|a - b|$, n'est pas perçue quand elle est inférieure ou égale à une certaine valeur $i$, et les points $a$ et $b$ sont donc indiscernables. On peut donc définir ainsi la relation d'indiscernabilité :

(7.41) $v_i(a \approx b) = \mathbf{V}$ si et seulement si $|a - b| \leq i$.

L'indiscernabilité, représentée par $\approx$, n'est manifestement pas transitive. On *n*'a donc *pas*

(7.42) $a \approx b, b \approx c \models a \approx c$,

quoique l'on ait

(7.43) $a \approx b \models b \approx c \rightarrow a \approx c$.

Cette différence entre 7.42 et 7.43 *ne* peut *pas* être exprimée par le conditionnel philonien car pour ce conditionnel l'équivalence suivante est vérifiée :

(7.44) $P, Q \models R$ si et seulement si $P \models Q \supset R$.

Pour $\rightarrow$, au contraire, on a seulement l'implication dans un sens :

(7.45) Si $P, Q \models R$ alors $P \models Q \rightarrow R$.

### 7.7.6. *Retour à la dérivabilité*

Les trois axiomes de la relation de conséquence logique s'appliquent aussi à la relation syntaxique de dérivabilité :

**Réflexivité** $(R)$

(7.46) $X \vdash X$ ,qui se lit "de $X$ on peut déduire $X$".$A \vdash A$.

**Monotonie** $(M)$[17]

(7.47) $$\frac{X \vdash Y}{X, X' \vdash Y, Y'}$$

où la virgule s'emploie comme dans les séquents.

**Transitivité** $(T)$

(7.48) $$\frac{X \vdash Y \quad Y \vdash Z}{X \vdash Z}$$

La notion de dérivabilité représentée par le signe $\vdash$ et définie par l'axiome $R$ et les deux règles d'inférence $M$ et $T$ est *logiquement antérieure* aux règles qui gouvernent les connecteurs. Une fois qu'on a caractérisé le sens de la relation de dérivabilité, on peut, au moyen de cette relation, introduire des connecteurs ou montrer que certains connecteurs *ne peuvent pas* être introduits.

Par exemple, on peut introduire et définir indirectement le connecteur "et" $(\wedge)$ au moyen de la règle d'*introduction*

(7.49) $A, B \vdash A \wedge B$,

et des règles d'*élimination*

(7.50) $A \wedge B \vdash A$,

(7.51) $A \wedge B \vdash B$.

Supposons à présent, avec Prior, que nous voulions ériger le mot "tonk" en connecteur au moyen de la règle d'*introduction*

(7.52) $A \vdash A$ tonk $B$,

qu'on peut lire "de $A$ on peut déduire $A$ tonk $B$", et de la règle d'*élimination*

(7.53) $A$ tonk $B \vdash B$,

[17] Des logiques non-monotones qui renoncent à la deuxième propriété ont été construites pour formaliser le raisonnement révisable. Voir à ce sujet Léa Sombé, 1989; M. Reinfrank, J. de Kleer, M. L. Ginsberg, 1989; Ph. Smets et al., 1988; A. Thayse et al., 1988, 1989.

qu'on peut lire "de $A$ tonk $B$ on peut déduire $B$".

On voit tout de suite qu'en raison de la transitivité de la relation de dérivabilité, l'application de la règle d'introduction du prétendu connecteur *tonk* suivie de l'application de sa règle d'élimination conduit à la conclusion $A \vdash B$, c'est-à-dire à la conclusion que n'importe quoi est dérivable de n'importe quoi. Ce pseudo-connecteur efface complètement la distinction entre ce qui est dérivable et ce qui ne l'est pas. Avec un tel connecteur, *tout* devient dérivable. Si l'on disposait déjà de la négation, on arrive immédiatement à une contradiction : il suffit de substituer, dans $A \vdash B$, $A$ à $B$ et ensuite $\neg A$ à $B$. On obtient alors $A \vdash A$ et $A \vdash \neg A$, en d'autres termes $A \vdash A, \neg A$.

Il est clair que "tonk" n'est pas un connecteur admissible. On notera ici l'importance d'une connaissance préalable de la notion de dérivabilité. Comme l'écrit Belnap,

> [...] nous ne définissons pas nos connecteurs *ab initio*, mais plutôt en termes d'un *contexte de dérivabilité donné préalablement* (Belnap,1962, 1967, p. 133).

L'extension du langage obtenue par l'introduction d'un nouveau connecteur doit être une *extension conservative*, en ce sens que l'ensemble des énoncés de dérivabilité (énoncés contenant $\vdash$) d'où le nouveau connecteur est absent, ne peut pas s'accroître par suite de l'introduction du nouveau connecteur. Cette exigence a été violée dans le cas de "tonk" puisqu'à l'aide de ce connecteur on obtient un énoncé de dérivabilité ne contenant pas "tonk", $(A \vdash B)$, énoncé qu'on ne pouvait pas obtenir avant d'avoir introduit "tonk" dans le langage.

Une cinquième et dernière contrainte générale qui pèse sur la relation de dérivabilité est l'obligation pour celle-ci d'être fermée sur l'opération de substitution. Czelakowski et Malinowski formulent ainsi cette exigence :

(7.54) Si $X \vdash Y$, alors $\sigma X \vdash \sigma Y$,

pour toute substitution $\sigma$ de signes du langage $\mathcal{L}$ qui appartiennent à la catégorie syntaxique requise[18]. Lorsque la relation de dérivabilité satisfait cette exigence, elle est dite *structurale*.

On vient de voir que la relation de *dérivabilité* doit être invariante par rapport à la *substitution*. Précédemment, on a vu que la relation de *conséquence logique* était invariante par rapport à l'*interprétation* ($\Gamma \models \Delta$ si toutes les interprétations qui rendent vraies les propositions membres de l'ensemble $\Gamma$ rendent vraie au moins une proposition de l'ensemble $\Delta$). On notera que la relation $\models$ unit un *ensemble* $\Gamma$ de prémisses à un *ensemble* $\Delta$ de conclusions.

Quelle est la portée de cette double exigence ? Nous allons voir qu'elle nous met en mesure d'expliquer le caractère contraignant du raisonnement logique.

[18] On ne peut évidemment pas substituer une variable individuelle ou une constante individuelle à une variable propositionnelle.

## 7.8. La nécessité du raisonnement logique

On se souvient de la définition aristotélicienne du syllogisme :

> Le *syllogisme* est un discours dans lequel, certaines choses étant posées, quelque chose d'autre que ces données en découle *nécessairement* par le seul fait de ces données.

De quelle sorte de nécessité s'agit-il ?
G. G. Granger répond :

> Le syllogisme en tant que tel ne fonde que la nécessité conditionnelle de sa conclusion, la vérité des prémisses n'étant qu'assurée (G. G. Granger, 1976, p. 11)

et insiste "sur le caractère formel du lieu syllogistique" (Vuillemin, 1984, p. 301). Le développement qui va suivre s'inscrit dans le prolongement de cette réponse et l'étend au raisonnement logique en général.

Dans son *Initiation aux mathématiques nouvelles à partir de problèmes techniques* (1970), P.V. Grosjean observe que l'implication en logique bivalente est une relation d'ordre total non-strict qui ordonne l'ensemble des *valeurs de vérité* $\{0,1\}$. Remplaçant le signe $\supset$ par le signe $\leq$, on a en effet[19]

(7.55) $\quad 0 \leq 0 \quad 0 \leq 1 \quad 1 \leq 1 \quad (\star).$

Mais l'implication n'ordonne pas seulement les valeurs de vérité,

> [Elle] ordonne aussi les propositions, créant des chaînes de propositions, chaînes appelées des *raisonnements*. Si toutes les implications sont vraies (raisonnement correct), et si la première proposition est vraie, il en va de même pour toutes les propositions conséquentes : le vrai implique le vrai et n'implique que le vrai, ainsi que le résume l'inégalité $1 \leq 1$, laquelle n'est qu'une réflexivité. Mais si la première proposition est fausse, on ne peut rien conclure d'un raisonnement : le faux implique n'importe quoi, puisque "faux" est ordonné, par rapport à "faux" et à "vrai", selon $0 \leq 0$, $0 \leq 1$ (Grosjean, 1970, p. 92).

L'explication de P.V. Grosjean rend compte du caractère contraignant de la *relation de conséquence* qui lie les vérités entre elles, relation que le logicien Bolzano appelle *Abfolge* dans sa *Wissenschftslehre* (1837), mais il laisse dans l'ombre une autre relation de conséquence, celle qui peut lier aussi des propositions fausses. Or cette deuxième relation existe bel et bien; par exemple, nous voulons pouvoir dire que certaines propositions découlent *nécessairement* du géocentrisme, bien que celui-ci soit faux. Bolzano a fait place à cette autre relation de conséquence, reliant cette fois des propositions aussi bien fausses

[19] Dans les formules 7.55, nous employons exceptionnellement "0" et "1" pour désigner les valeurs de vérité **F** et **V**, étant donné que l'auteur précité les traite comme les nombres 0 et 1.

que vraies. Il l'appelle *Ableitbarkeit* — littéralement "déductibilité" —, mais il désigne par ce mot, en fait, une relation sémantique comme celle que l'on représente aujourd'hui par $\models$.

Bolzano, citant la définition du syllogisme que nous avons rappelée, la commente en ces termes :

> puisqu'il ne fait pas de doute qu'Aristote admettait que la relation de déductibilité peut aussi relier des propositions fausses, le "découle nécessairement" ne peut guère être interprété autrement que comme ceci : que la conclusion devient vraie *chaque fois* que les prémisses sont vraies (Bolzano, 1837, 1972, pp. 219–220).

Parler de propositions qui *deviennent* vraies quand d'autres propositions le sont suggère que certaines propositions changent de valeur de vérité et sont tantôt fausses, tantôt vraies. Mais Bolzano écarte cette interprétation :

> Or il est évident que nous ne pouvons dire d'une et même classe de propositions que l'une [la conclusion] devient vraie *chaque fois* que les autres [les prémisses] sont vraies, sauf si nous traitons certaines de leurs parties comme variables. En effet, des propositions dont aucune partie ne change ne sont pas tantôt vraies, tantôt fausses, elles sont toujours soit vraies soit fausses. Quand nous disions de certaines propositions que l'une devenait vraie aussitôt que les autres l'étaient, nous n'avions pas en vue ces propositions elles-mêmes, mais les propositions en nombre infini qui peuvent être générées à partir d'elles si certaines de leurs idées sont remplacées par d'autres arbitrairement choisies (Bolzano, *Ibid.*, p. 220).

Ce ne sont donc pas les valeurs de vérité des propositions qui varient, mais les constituants de ces propositions.

La question qui se posait, c'était de savoir en quel sens une proposition fausse comme, par exemple "Tous les singes sont des oiseaux", peut *découler nécessairement* des propositions fausses "Tous les poissons sont des oiseaux" et "Tous les singes sont des poissons". La réponse de Bolzano, dans l'exemple considéré, serait que toutes les propositions de la forme "Tous les $S$ sont $P$" découlent nécessairement des propositions "Tous les $M$ sont $P$" et "Tous les $S$ sont $M$" en ce sens que *toutes* les substitutions de termes syntaxiquement admissibles aux variables $S$, $M$ et $P$ qui transforment les deux prémisses en propositions vraies, transforment aussi, *ipso facto*, la conclusion en proposition vraie.

La définition de Bolzano préfigure celle que donne Tarski en 1936, définition qui fait appel au concept de *modèle* défini en ces termes :

> Soit une classe quelconque de propositions $L$. Nous remplaçons toutes les constantes extra-logiques que l'on trouve dans les propositions de $L$ par les variables correspondantes, les mêmes constantes étant remplacées par les mêmes variables, les constantes distinctes étant remplacées par des variables distinctes. De cette façon, nous obtenons une classe $L'$ de

> fonctions propositionnelles. Toute suite quelconque d'objets qui satisfait toutes les fonctions propositionnelles de la classe $L'$ sera appelée *modèle*, ou *réalisation de la classe de propositions* $L$ [...] (Tarski, 1936, 1974, p. 149)

Armé de cette définition préalable, Tarski définit la conséquence logique :

> La proposition $X$ *suit logiquement* des propositions de la classe $K$ si et seulement si tout modèle de la classe $K$ est aussi un modèle de la proposition $X$ (Tarski, *Ibid.*, p. 150).

La définition traditionnelle de la locution "raisonnement logiquement valide" fait appel, comme la définition aristotélicienne du syllogisme, aux termes modaux "nécessairement" ou "impossible". Un raisonnement est logiquement valide s'il est *impossible* que ses prémisses soient vraies et sa conclusion fausse. Hélas la notion d'impossibilité a plusieurs sens : il y a des impossibilités mathématiques — la quadrature du cercle —, des impossibilités physiques — le mouvement perpétuel —, des impossibilités juridiques — la bigamie en droit français.

Le progrès décisif réalisé par Bolzano et Tarski, c'est d'avoir formulé une définition de la validité du raisonnement qui ne fait plus appel à des notions modales : un raisonnement est valide si et seulement si *toutes* les *substitutions* ou *interprétations* qui transforment les prémisses en propositions vraies transforment la conclusion en proposition vraie.

La définition de Bolzano, comme celle de Tarski, exige que l'on commence par diviser le vocabulaire en deux parties : une partie composée de *termes extra-logiques* sur lesquels on opérera des substitutions ou des interprétations, et une partie composée de *termes logiques* qui resteront à l'abri de la substitution ou de l'interprétation. Si on ne faisait pas cette distinction et si on s'autorisait à substituer ou à interpréter librement *toutes* les expressions d'un raisonnement, aucun raisonnement formellement valide ne subsisterait. Reprenons l'exemple cité plus haut : "Tous les poissons sont des oiseaux et tous les singes sont des poissons donc tous les singes sont des oiseaux". Ce raisonnement est valide parce que toutes les substitutions uniformes opérées sur les noms communs "poissons", "oiseaux" et "singes" qui transforment les prémisses en propositions vraies, transforment en même temps les conclusions en propositions vraies, mais il aurait cessé d'être valide si au lieu de faire des substitutions sur les noms communs, on avait également fait des substitutions sur les quantificateurs et que l'on avait remplacé, par exemple, les occurrences de "tous", par des occurrences de "quelques". En effet, nous obtenons alors un raisonnement de la forme

(7.56) Quelques $M$ sont $P$,
Quelques $S$ sont $M$,
Donc quelques $S$ sont $P$,

dont il est facile de montrer par un exemple qu'il n'est pas valide :

(7.57) Quelques mausolées sont des cubes,
Quelques pyramides sont des mausolées,
Donc quelques pyramides sont des cubes.

Les deux prémisses de ce raisonnement sont vraies, mais sa conclusion est fausse.

La notion de validité d'un raisonnement est donc *relative* au lexique que l'on se donne au départ, c'est-à-dire à l'ensemble des expressions sur lesquelles on ne fera pas de *substitution* et que l'on prendra dans leur *interprétation standard*. *Tous* les termes logiques rencontrés jusqu'à présent (les connecteurs, les quantificateurs, le prédicat d'identité) sont unanimement acceptés comme termes logiques. En revanche, il n'y a pas d'accord sur la manière d'étendre la liste : Quine conteste, par exemple, le droit de parler de logique épistémique[20]. La séparation entre termes logiques et termes non logiques semble assez floue et arbitraire.

### 7.8.1. *La conception formelle de la validité*

La reconnaissance du caractère *flou* et *arbitraire* de la division du lexique en termes logiques et termes extra-logiques met-elle en péril l'explication du caractère contraignant de la validité logique donnée par Bolzano et Tarski ? On pourrait le craindre puisque cette explication repose sur la distinction entre *forme* et *contenu*, la forme étant le résidu que l'on obtient quand on a remplacé les constantes extra-logiques par des variables.

Nous avons dit des définitions de la conséquence proposée par Bolzano et Tarski qu'elles avaient le mérite de rendre compte du caractère contraignant du raisonnement logique, tout en éliminant la notion modale de nécessité qui figure dans l'expression aristotélicienne "découle nécessairement", notion qu'elles remplacent par la notion extensionnelle d'universalité exprimée par le quantificateur ("toutes les substitutions", "tous les modèles"). Tarski lui-même dit explicitement que la vérité de la prémisse *contraint* la conclusion à être vraie :

> En particulier, on peut démontrer à partir de cette définition (définition en termes de modèles), que toute conséquence de propositions vraies *doit* être vraie (Tarski, *Op. cit.*, p. 150. C'est nous qui soulignons.)

Dans un article récent (1983), Etchemendy conteste l'affirmation de Tarski. Selon Etchemendy, la définition du raisonnement valide, qui est le corollaire de la définition de conséquence proposée par Bolzano et Tarski, permet seulement de prouver que si un raisonnement est valide et que ses prémisses sont vraies, alors sa conclusion *est* vraie, elle ne permet pas de prouver l'affirmation plus

[20] La logique épistémique étudie les raisonnements formels dans lesquels "*a* sait que *p*" est rangé parmi les termes du lexique logique. Par exemple, le schéma d'inférence suivant relève de la logique épistémique : "*a* sait que *p* donc *p*" (voir tome II).

ambitieuse de Tarski selon laquelle lorsqu'un raisonnement est valide et que ses prémisses sont vraies, la conclusion *doit* être vraie.

Etchemendy considère le raisonnement suivant dont la prémisse et la conclusion sont chacune empiriquement vraies. Les énoncés sont écrits en 1983, à un moment où Ronald Reagan occupe la présidence des Etats-Unis et où Etchemendy est dans le New Jersey.

(7.58) Reagan est président,
donc Etchemendy est dans le New Jersey.

Puisque la distinction entre termes logiques et non logiques est arbitraire, convenons d'inclure dans le lexique logique tous les mots figurant dans ce raisonnement. Cette décision suffit à ériger le raisonnement ci-dessus en raisonnement *valide*. En effet, toutes les substitutions qui rendent la prémisse vraie rendent la conclusion vraie, puisque la seule substitution autorisée est la substitution nulle qui laisse tout en place. Nous nous trouvons donc devant un raisonnement valide dont la prémisse est vraie, mais où la conjonction de l'affirmation de la prémisse et de la négation de la conclusion est simplement fausse, mais pas impossible.

Le raisonnement d'Etchemendy est irréprochable, mais est-il vraiment pertinent ? Tarski avait, en effet, envisagé les cas de ce type. Parlant de l'opposition entre termes logiques et non logiques, il écrit :

> [...] je ne connais aucune base objective qui nous permette de dresser une frontière nette entre ces deux groupes de termes. Il me semble possible d'inclure parmi les termes logiques quelques termes qui sont habituellement considérés comme extra-logiques, sans pour autant obtenir des conséquences qui s'opposeraient fortement à l'usage ordinaire. A la limite, nous pourrions considérer tous les termes du langage comme logiques. Le concept de conséquence *formelle* coïnciderait alors avec celui de conséquence *matérielle*. La proposition $X$ se déduirait alors de la classe de propositions $K$ si $X$ était vraie, ou si au moins une des propositions de la classe $K$ était fausse (Tarski, *Op. cit.*, p. 151).

Le contre-exemple d'Etchemendy est destiné à réfuter l'affirmation selon laquelle un raisonnement formellement valide contraint, par sa forme même, la conclusion à être vraie si les prémisses sont vraies. Il nous semble que ce contre-exemple n'est pas pertinent dans la mesure où il *n'illustre pas* la relation de conséquence formelle, mais seulement la relation de conséquence matérielle correspondant au conditionnel philonien. Nous pouvons donc continuer à *localiser la source du caractère contraignant du raisonnement logique* dans sa forme et à remplacer dans la définition de la validité l'idée de *nécessité* par l'idée *d'universalité*.

## 7.9. Vérité et preuve

En guise de transition entre ce chapitre consacré à la sémantique, notamment à la notion de vérité, et le chapitre suivant consacré à la notion de preuve, nous étudierons ici la relation entre vérité et preuve.

### 7.9.1. *Conditions d'adéquation*

Si notre ambition se limitait à définir la vérité pour un nombre fini de phrases, par exemple pour les phrases membres de l'ensemble $\Gamma$

(7.59) $\Gamma$ = { La neige est blanche, 4 est pair },

il nous suffirait d'écrire :

(7.60) "La neige est blanche" est vrai en français[21] si et seulement si la neige est blanche.

(7.61) "4 est pair" est vrai en français si et seulement si 4 est pair.

Les propositions 7.60 et 7.61 sont des *définitions partielles* de la vérité. Ce sont des définitions — le terme "vrai" est absent du définissant — mais partielles parce qu'elles ne définissent la vérité que pour une phrase singulière.

Or, il est exclu de généraliser ces définitions partielles de manière expéditive et d'écrire :

"$p$" est vrai si et seulement si $p$.

En effet, l'expression

(7.62) "$p$"

est une constante individuelle et la lettre mise entre guillemets

(7.63) $p$

n'est pas plus une *variable propositionnelle* que l'$x$ de "xénophobe" n'est une *variable individuelle*.

Pour exprimer ce qu'il y a de commun à 7.60 et à 7.61, on doit recourir au schéma

(7.64) $X$ est vrai si et seulement si $p$

où $X$ est une variable qui prend comme substitut le nom de la phrase écrite à droite de la conjonction *si et seulement si* et comme valeur la phrase elle-même.

Tarski impose deux *conditions d'adéquation* à une définition de la vérité; conditions qui, notons-le, sont satifaites par la définition en termes de satisfaction par toutes les suites.

---

[21] Il faut préciser de quelle langage-objet fait partie la phrase dont on énonce les conditions de vérité. En effet, une même phrase peut avoir des sens différents selon la langue à laquelle elle appartient. On donne souvent l'exemple amusant d'une phrase en écriture *phonétique* qui, selon qu'elle est entendue comme phrase latine ou comme phrase française, admet deux sens différents (*Hector apportavit legato aurum*).

1. Une condition d'*adéquation matérielle* : impliquer toutes les instances du schéma 7.64 (ce qui ne nous assure pas que nous avons bien cerné l'*extension* de "vrai").
2. Une condition d'*adéquation formelle* : écarter le paradoxe du menteur.

### 7.9.2. *Le paradoxe du menteur*

Voici le *paradoxe du menteur*, dont la première version est attribuée au logicien grec Eubulide, paradoxe qui causa la mort prématurée de Philetas de Cos. Une définition satisfaisante de la vérité, c'est-à-dire qui respecte au minimum l'*extension* du mot "vrai", doit, affirme Tarski, impliquer toutes les instances du schéma

(7.65) $X^{(p)}$ est vrai si et seulement si $p$,

où $p$ est une variable qui prend comme substituts des *phrases*, et où $X^{(p)}$ est un nom de la phrase $p$. L'exposant dont Tarski se passe n'est là que pour mémoire.

Considérons à présent la phrase auto-référentielle

(7.66) La phrase 7.66 est fausse.

Si nous instancions le schéma 7.65 à l'aide de 7.66, nous obtenons

(7.67) "la phrase 7.66 est fausse" est vraie si et seulement si $\underbrace{\text{la phrase 7.66}}_{a}$ est fausse. Or l'inspection des lignes qui précèdent nous apprend que l'affirmation d'identité $(a = b)$ suivante est vraie :

(7.68) $\underbrace{\text{la phrase 7.66}}_{a} = \underbrace{\text{“la phrase 7.66 est fausse”}}_{b}$.

En vertu de la loi de substituabilité des identiques (sections 1.5.2 et 1.5.3), on peut remplacer $a$ par $b$ dans 7.67, et on obtient une contradiction de la forme

(7.69) $p \equiv \neg p$,

à savoir

(7.70) "La phrase 7.66 est fausse" est vraie si et seulement si "la phrase 7.66 est fausse" est fausse.

### 7.9.3. *Solution du paradoxe du menteur*

Dans la langue naturelle, nous pouvons former des phrases qui affirment leur propre fausseté, comme c'est le cas de la phrase 7.66. Comme la langue formelle est une langue artificielle qui aura toutes les bonnes propriétés qu'il nous plaira de lui donner, on peut *décréter* que la langue formelle est hiérarchisée en langue-objet, métalangue, métamétalangue... (cf. section 2.2.2) et bannir comme mal

formées les phrases de la langue-objet qui contiendraient le mot "faux" ou le mot "vrai", comme c'est le cas de 7.66. Ce décret n'empêche nullement de définir le prédicat "vrai" et de formuler le schéma 7.65. Il ne fait rien de plus que déplacer cette définition et ce schéma en les situant dans le *métalangage*, lequel doit être plus riche que la langue-objet, en cela qu'il contient toutes les expressions de la langue-objet et, par surcroît, les *noms* de ces expressions et de prédicats tels que "vrai", "faux", "satisfait", etc.

### 7.9.4. *Métalangage sémantique et métalangage syntaxique*

Nous avons vu, en 2.4.1, qu'il est possible de formuler dans le langage-objet de l'arithmétique les expressions du métalangage syntaxique, telles que, par exemple :

(7.71) "$2 + 2 = 4$" est démontrable,

(7.72) "$0 + 0 = 0,\ \exists x(0 + 0 = x)$" est une inférence,

grâce à la codification gödélienne.

Dans une théorie assez riche pour contenir le langage de l'arithmétique élémentaire (c'est-à-dire "0", "successeur de" — et du même coup le nom de tous les entiers naturels construits récursivement à partir de "0" et de "successeur de" —, "nombre", les symboles "+", "×", et le prédicat "="), il est possible de représenter par des nombres gödéliens — donc par des nombres naturels — les phrases. Il est possible aussi de représenter les prédicats du métalangage syntaxique "démontrable", "réfutable"... par des prédicats arithmétiques (définis à l'aide de termes de l'arithmétique). En revanche, il n'est pas possible, par la même technique, d'incorporer au langage-objet de l'arithmétique les prédicats du métalangage sémantique "vrai", "faux", "satisfait"... car une telle incorporation rendrait inconsistante la théorie qui contiendrait la définition de "vrai" formulée dans le langage-objet de l'arithmétique. Elle permettrait, en effet, d'y générer le paradoxe du menteur. Il y a donc une différence cruciale entre les deux phrases auto-référentielles suivantes :

(7.73) la phrase 7.73 est fausse,

(7.74) la phrase 7.74 est indémontrable.

La première génère un paradoxe, la seconde ne le fait pas.

### 7.9.5. *Ecart entre les notions de vérité et de démontrabilité*

En exploitant cette différence entre le prédicat "vrai" et les prédicats "démontrable", "réfutable"..., on retrouve par une autre voie le célèbre théorème démontré par Gödel en 1931, que Tarski énonce dans la dernière phrase du passage suivant :

[...] l'ensemble des nombres prouvables [c'est-à-dire des nombres gödéliens de théorèmes de l'arithmétique] ne coïncide pas avec l'ensemble des nombres vrais [c'est-à-dire des nombres gödéliens de phrases vraies], puisque le premier ensemble est définissable dans le langage de l'arithmétique tandis que le second ne l'est pas. En conséquence, l'ensemble des phrases prouvables ne coïncide pas non plus avec l'ensemble des phrases vraies. D'autre part, en utilisant la définition de la vérité, nous montrons aisément que tous les axiomes de l'arithmétique sont vrais, et que toutes les règles de preuve sont infaillibles. Ainsi, toutes les phrases prouvables sont vraies, donc la réciproque n'a pas lieu [*does not hold*]. Notre conclusion finale est : il existe des phrases formulées dans le langage de l'arithmétique mais qui ne peuvent être prouvées formellement sur base des axiomes et des règles de preuve acceptées en arithmétique (Tarski, 1969, pp. 298–299).

# Chapitre 8
# Théorie de la preuve

## 8.1. La crise des fondements des mathématiques

En 1902, Russell a montré que le principe d'abstraction

(8.1) $\exists y \forall x (x \in y \equiv Fx)$,

qui se lit

(8.2) Il existe un ensemble qui a pour membres tous les êtres qui satisfont le prédicat $F$ et eux seuls

conduit à une contradiction de la forme

(8.3) $p \equiv \neg p$

lorsqu'on prend comme prédicat $F$ le prédicat $x \notin x$.

Ce paradoxe surprend. De prime abord, on serait tenté de dire qu'il y a des classes qui sont membres d'elles-mêmes (la classes des classes est dans ce cas puisqu'elle est une classe) et des classes qui ne sont pas membres d'elles-mêmes (la classe des chaises, par exemple, n'est pas membre de la classe des chaises, puisqu'elle n'est pas une chaise). On penserait donc qu'on ne risque rien à affirmer qu'il existe une classe dont sont membres toutes et seulement les classes $x$ qui ne sont pas membres d'elles-mêmes.

Russell proposa une première solution, qui consiste à modifier le *langage* de la théorie naïve des ensembles, en introduisant dans les règles de formation des contraintes qui ont pour effet d'écarter comme expression mal formée l'assemblage $x \notin x$. Au lieu d'adopter un langage qui ne comporte qu'une espèce de variables, on adoptera un langage qui "contient une hiérarchie dénombrable de *niveaux* de variables" (Frankel, Bar-Hillel, Levy, 1973, p. 158) :

> Chaque variable appartient à un et un seul niveau seulement. Les *numéros de niveau* $1, 2, 3 \ldots$ seront indiqués par des chiffres en exposants à droite. Les seules formules atomiques de [la théorie des types] $T$ seront [...] des formules de la forme "$\ldots \in \_\_$", et de la forme "$\ldots = \_\_$", mais désormais le numéro de niveau de la variable de gauche des formules de la première espèce doit être inférieur d'une unité exactement au nombre de niveau de la variable de droite; en d'autres termes, $x^i \in y^j$ est une formule si et seulement si $j = i + 1$[...] (Frankel, Bar-Hillel, Levy, *Ibid.*)

Au lieu de modifier le *langage*, on peut modifier la *théorie* elle-même, et remplacer le principe litigieux par un principe plus restreint, c'est ce qu'a fait Zermelo en 1908. Aujourd'hui, ce principe plus restreint, appelé *Aussonderung Axiom* (axiome de séparation), est incorporé à la théorie des ensembles la plus courante, celle de Zermelo-Frankel (1922). Il est formulé ainsi :

(8.4) $\quad \forall z \exists y \forall x (x \in y \equiv (x \in z \wedge \Phi x))$.

A toute fonction propositionnelle correspond un ensemble, pour autant que l'on ait déjà un ensemble *donné* au préalable. Ce n'est que si cette condition est remplie que la fonction propositionnelle génère un ensemble, qui sera un sous-ensemble de l'ensemble donné. Si l'on adopte cette solution, on ne doit plus rejeter $x \notin x$ parmi les expressions mal formées[1].

Le mathématicien hollandais Brouwer, dont la philosophie s'inspire de la philosophie de Kant, est beaucoup plus critique à l'égard de la notion d'ensemble. Pour Brouwer,

> la locution "l'ensemble de tous les..." ne renvoie pas à une entité constructible et doit recevoir un contenu [de sens] nouveau. En effet, à la notion classique d'ensemble correspondent deux notions intuitionnistes, celle de déploiement et celle d'espèce — un déploiement étant défini par un mode commun de génération de ses éléments (constructibles), et une espèce étant définie par la propriété caractéristique qui peut être assignée aux entités mathématiques qui ont été ou auraient pu être construites *avant* qu'on ne définisse l'espèce (Körner, 1960, p. 128).

Hilbert a adopté une troisième position qui, dans une certaine mesure, concilie le souci de rigueur des intuitionnistes avec la volonté de conserver l'acquis des mathématiques classiques. Hilbert distingue la mathématique de la métamathématique. La seconde est une discipline qui prend pour objet non pas les objets mathématiques (nombres, grandeurs...), mais la mathématique elle-même, à savoir la science mathématique coulée en systèmes d'axiomes. Le

[1] La raison pour laquelle la présence de cet axiome écarte le paradoxe de Russell est la suivante. Etant donné n'importe quel ensemble A, si dans l'axiome de séparation nous substituons $x \notin x$ à $\Phi x$, nous obtenons l'ensemble $R = \{x : x \in A \wedge x \notin x\}$ où $y \in R \equiv y \in A \wedge y \notin y$. Instanciant $y$ en $R$ nous obtenons $R \in R \equiv R \in A \wedge R \notin R$ qui n'est pas une contradiction mais qui peut se simplifier en $R \notin R \wedge R \notin A$ (cf volume II et Copi, 5e éd., p. 180).

programme que s'assigne Hilbert est d'établir la non-contradiction de théories mathématiques existantes telles que, par exemple, la théorie axiomatique des nombres. L'école de Hilbert n'est pas opposée à l'emploi par le mathématicien de notions non constructives, telles que la notion d'ensemble, elle n'est opposée à l'usage de ces notions qu' en *métamathématique*.

Il n'a pas fallu attendre Hilbert pour savoir qu'une démonstration n'est acceptable que si elle est finie, ni pour savoir qu'une régression à l'infini est un raisonnement vicieux. La contrainte finitiste imposée par l'école de Hilbert est quelque chose de plus subtil. Sans entrer dans les détails, on peut donner une idée de la différence en comparant le raisonnement par récurrence admissible en *métamathématique* au raisonnement par récurrence admissible en *mathématique*. On trouve le premier à l'oeuvre, par exemple, dans une démonstration dans laquelle,

> pour montrer [...] que toute proposition vraie dans une théorie possède une propriété $A$, nous montrerons successivement que les propositions que l'on admet comme vraies au début de la théorie possèdent la propriété $A$, et que, si des propositions vraies possèdent la propriété $A$, il en est de même pour celles que l'on peut en déduire en appliquant une des règles du raisonnement (Herbrand, 1930, 1986, p. 38).

Cette récurrence métamathématique est, en quelque sorte, *opératoire* et *procédurale* :

> Ici [en métamathématique], [la manière d'employer le raisonnement par récurrence] n'est jamais que l'indication, en une seule formule, d'un procédé qu'il faudra appliquer un certain nombre de fois dans chaque cas particulier (Herbrand, 1930, 1986, p. 39).

Elle diffère profondément de la récurrence mathématique dans laquelle peuvent intervenir des *éléments idéaux* :

> En mathématique, au contraire, il arrive que l'on utilise ce raisonnement dans des cas où l'on considère des concepts qui n'ont pas de représentation matérielle possible, au contraire de nos signes, comme par exemple l'ensemble des entiers ou des nombres réels (Herbrand, 1930, 1986, p. 39).

Alors que Hilbert n'imposait l'exigence finitiste qu'au logicien engagé dans une preuve de non-contradiction, Herbrand, "plus royaliste que le roi" comme l'écrit J. van Heijenoort (1982, p. 77), étend cette exigence à toute recherche logique, ce qui le conduit à mettre en question les notions de validité ou de consistance dans un domaine infini.

## 8.2. Les théorèmes de Löwenheim et de Skolem

A première vue, plus le domaine d'individus s'accroît, plus nombreux sont les schémas quantificationnels [2] qui peuvent être satisfaits simultanément (c'est-à-dire transformés en propositions vraies). Par exemple, dans un univers constitué *d'un seul* individu, le schéma quantificationnel

(8.5) $\exists x Fx \wedge \exists x \neg Fx$

ne peut être satisfait, mais il peut l'être dans un domaine à *deux* individus au moins.

De même, le schéma

(8.6) $\exists x(Fx \wedge Gx) \wedge \exists x(Fx \wedge \neg Gx) \wedge \exists x(\neg Fx \wedge Gx)$

ne peut être satisfait dans un domaine à *deux* individus mais il peut l'être dans un domaine à *trois* individus au moins. Il y a aussi des schémas quantificationnels qui ne sont sémantiquement consistants dans aucun domaine numérique fini, mais qui le sont dans le domaine des entiers naturels. C'est le cas, par exemple, du schéma qui résulte de la conjonction des trois formules (8.7) du calcul des prédicats, si l'on interprète le prédicat à deux places $R$ comme exprimant la relation *plus petit que* (Smullyan, 1968, p. 63).

(8.7) $\forall x \exists y Rxy,$
$\neg \exists x Rxx,$
$\forall x \forall y \forall z(Rxy \wedge Ryz \supset Rxz).$

On s'attendrait dès lors à ce que le choix d'un domaine infini qui soit de cardinalité supérieure à celle des entiers naturels, un domaine qui est un *ensemble infini indénombrable* tel que le domaine constitué par l'ensemble des nombres réels, permette d'augmenter encore le nombre des schémas du calcul des prédicats qui peuvent être satisfaits simultanément dans ce modèle. Le théorème de Löwenheim a démenti cette attente.

En 1915, Löwenheim a établi le théorème 8.8 :

(8.8) Un schéma quantificationnel $S$ est sémantiquement consistant dans une interprétation et un domaine non vide si et seulement si il est vrai pour une interprétation dans l'ensemble des entiers positifs.

Etant donné les rapports existant entre consistance sémantique et validité, on peut reformuler ainsi le théorème de Löwenheim :

[2] Pour aligner notre terminologie sur celle des auteurs cités, nous appelons ici schémas les formules du calcul des prédicats dont les constantes prédicatives ($F$, $G$, $R$...) n'ont pas été interprétées. Un schéma est *valide*, *sémantiquement consistant* ou *sémantiquement inconsistant* selon qu'il est transformé en proposition vraie par *toute*, par *quelque* ou par *aucune* interprétation de ces constantes dans un domaine non vide.

(8.9) Un schéma quantificationnel est valide, c'est-à-dire vrai dans toutes les interprétations dans un domaine non vide, si et seulement si il est vrai dans toutes les interprétations dans l'ensemble des entiers positifs,

ce qui peut s'écrire :

(8.10) $\models S$ si et seulement si $\|S\|_{\mathcal{I}} = \mathbf{V}^1$ dans $\mathcal{D} = \mathbf{N}^+$ pour toute interprétation $\mathcal{I}$.

Le théorème de Löwenheim est indiscutablement un théorème de métamathématique. Si on entend respecter les consignes de Hilbert, on devrait donc renoncer, pour la démonstration de ce théorème, aux notions non finitistes. Or c'est ce que ne fait pas Löwenheim, dont la démarche prête le flanc à cette critique de Herbrand :

> [...] il donne à la notion de "vérité dans un champ [domaine] infini" un sens intuitif, ce qui fait que sa démonstration n'atteint pas la rigueur qui nous semble désirable (Herbrand, 1930, 1968, p. 143),

et un peu plus loin :

> Nous pouvons dire que la démonstration de Löwenheim était suffisante en mathématique (Herbrand, *Ibid.*, p. 144),

ce qui sous-entend qu'elle ne l'est pas en *métamathématique*.

## 8.3. Le théorème de Herbrand

### 8.3.1. *Le passage de l'infini au fini chez Herbrand*

L'un des résultats majeurs de Herbrand peut s'énoncer comme suit :

(8.11) Etant donné une formule quelconque $F$ de la théorie de la quantification, il est possible d'engendrer de manière mécanique (et relativement simple) une suite de formules sans quantificateurs $F_0, F_1 \ldots$ telle que $F$ est démontrable en théorie de la quantification (c'est-à-dire dans l'un quelconque des systèmes équivalents qui cristallisent cette théorie) si et seulement s'il existe un nombre $k$ tel que $F_k$ est démontrable en logique propositionnelle ou, ce qui est équivalent, est valide par les tables de vérité.

On voit que l'on a ici un équivalent du théorème de Löwenheim dans lequel ne figurent que des notions finitistes : la notion ensembliste de validité a été remplacée par la notion syntaxique de démontrabilité et celle de vérité dans un ensemble infini par celle de vérité dans un ensemble fini.

Que Herbrand ait réussi à rabattre la théorie de la quantification, c'est-à-dire le calcul des prédicats du premier ordre dans toute sa généralité, sur le calcul des propositions, est surprenant.

On s'attendait à ce que l'entreprise fût impossible. Certes les schémas quantificationnels qui sont sémantiquement consistants dans un domaine fini ne posent pas de problème. On peut remplacer dans ces schémas les quantificateurs universels et existentiels par des conjonctions et des disjonctions finies qui ont autant de membres que le domaine contient d'individus. Mais que faire dans le cas des schémas qui ne sont sémantiquement consistants que dans un domaine infini ? N'y a-t-il pas un fossé infranchissable entre les deux types de situations ? A Herbrand revient le mérite d'avoir relevé ce défi et d'avoir jeté un pont entre le fini et l'infini grâce à l'idée de *suite infinie de domaines finis.*

Herbrand n'a donc pas accompli l'impossible tâche de réduire l'infini au fini. Il a fait autre chose. Il a remplacé la notion de domaine infini par celle de suite infinie de domaines finis, ces derniers pouvant être regardés comme des "approximations successives de ce domaine infini", selon l'heureuse expression de van Heijenoort (1982, p. 59).

S'il s'était arrêté là, Herbrand n'aurait pas vraiment satisfait l'exigence finitiste. En effet, l'infini n'a pas été exorcisé. Expulsé des domaines qui sont tous finis, il se réfugie dans la suite des domaines qui, elle, demeure infinie. N'est-ce pas une victoire à la Pyrrhus ? Nullement. Herbrand, en effet, a réussi à démontrer l'existence d'une asymétrie entre la consistance sémantique et l'inconsistance sémantique : pour prouver l'*inconsistance* d'un schéma quantificationnel, il suffit de considérer une *portion finie* de la suite infinie des domaines finis. Ceci nous allons l'établir en détail, ce qui demandera l'introduction d'un certain nombre de notions techniques.

### 8.3.2. *Forme clausale*

Nous avons établi en 6.4.2 que si $\Phi$ est la forme de Skolem d'une formule et que $\Gamma$ est l'ensemble des clauses obtenues en mettant $\Phi$ en forme clausale, $\Phi$ est inconsistant si et seulement si l'ensemble $\Gamma$ est inconsistant. Le problème qui se pose à présent est de démontrer que l'ensemble $\Gamma$ est inconsistant.

### 8.3.3. *Modèle de Herbrand*

Un ensemble $S$ de clauses est sémantiquement inconsistant si et seulement si la conjonction de ces clauses est fausse dans tous les modèles, c'est à dire pour toute interprétation et tout domaine. Il n'est ni possible ni nécessaire d'explorer tous les modèles. On peut se limiter à l'examen d'un type de modèle bien particulier, le modèle de Herbrand[3], $\mathcal{M} = \langle H, \mathcal{I} \rangle$ où $H$ est le *domaine de Herbrand* de l'ensemble de clauses $S$ et $\mathcal{I}_H$ l'*interprétation de Herbrand.*

[3]Les expressions "domaine de Herbrand" ("univers de Herbrand"), "interprétation de Herbrand" sont des expressions consacrées, contrairement à "modèle de Herbrand".

**Définition 8.1 : domaine de Herbrand.** Le domaine de Herbrand est stratifié en niveaux. Au niveau 0, il comprend l'ensemble $H_0$ constitué par :

(1) toutes les constantes d'individus figurant dans les clauses de $\alpha$. S'il y a de telles constantes, elles se trouvaient dans les formules avant la mise sous forme clausale ou elles sont issues de cette dernière transformation. S'il n'y en a aucune, on en crée une : $a$.

(2) tous les symboles de fonctions figurant dans les clauses de $\alpha$, que ces symboles figurent dans la formule dès avant la mise en forme clausale, ou qu'ils en soient issus. On leur donne comme arguments les constantes mentionnées en (1).

On obtient $H^{i+1}$ à partir de $H^i$ en formant l'union de $H^i$ et de l'ensemble de tous les symboles de fonction $f^n(t_1, \ldots, t_n)$ dans lesquels les $t_j$, $j = 1, \ldots, n$ sont membres de $H^i$.

On appelle domaine de Herbrand de $\alpha$ (d'après Chang et Lee, 1973, p.52), la réunion des ensembles $H_i$, pour $i \in \mathbf{N}$.

*Exemple.* Soit $\alpha = \{\exists z(\forall y((y > z) \supset \exists x(x^2 > 3)), \forall x \exists y(y = x^2))\}$

On commence par mettre les membres de $\alpha$ en forme clausale[4].

$\alpha' = \{\neg(y > a) \vee ((f(y))^2 > 3), g(x') = (x')^2\}$

$H_0 = \{a, 3, f(a), f(3), g(a), g(3), a^2, 3^2\}$

$H_1 = H_0 \cup \{f(f(a)), \ldots, f(a^2), \ldots, (f(a))^2, \ldots\}$

$H_2 = H_1 \cup \{f(f(f(a))), \ldots\}$

**Définition 8.2 : interprétation de Herbrand.** L'interprétation de Herbrand $\mathcal{I}_H$ applique chaque constante individuelle sur elle-même et applique chaque *symbole de fonction* "$f$" à $n$ places sur une fonction $h$ qui fait correspondre à $n$ termes $(t_1, \ldots, t_n)$ du domaine de Herbrand $H$ le terme $f(t_1, \ldots, t_n)$.

Dans l'exemple mentionné plus haut, on a $\mathcal{I}_H$ :

$$
\begin{array}{lll}
a \mapsto a & & \\
3 \mapsto 3 & & \\
(\ )^2 \mapsto h & : & a \mapsto (a)^2 \\
 & & 3 \mapsto (3)^2 \\
f(\ ) \mapsto h' & : & a \mapsto f(a) \\
 & & 3 \mapsto f(3) \\
etc \ldots & &
\end{array}
$$

[4] On pourrait craindre que le mélange de symboles mathématiques de fonctions telles que $(\_)^2$ et de fonctions de Skolem tels que $f(\_)$ n'engendre des non-sens. En effet, contrairement à un nombre, un symbole ne peut pas être élevé au carré. Cette crainte est vaine. On ne retiendra de $(\_)^2$ que son statut de foncteur unaire qui donne naissance à une constante individuelle quand on substitue une constante à sa place vide. Voir la définition suivante.

### 8.3.4. *Correspondance entre interprétations*

**Définition 8.4 : correspondance entre interprétations.** Soit $\mathcal{I}$ une interprétation quelconque dans un domaine $\mathcal{D}$, l'interprétation de Herbrand $\mathcal{I}_H$ pour un ensemble $S$ de formules est dite *correspondre* à $\mathcal{I}$ et sera noté $\mathcal{I}_H^*$ si les conditions suivantes sont satisfaites :

1. chaque élément $h_i \in H_S$ est appliqué sur un élément $d_i \in \mathcal{D}$.
2. si le prédicat $P(x_1, \ldots, x_n)$ est vrai (faux) du n-tuple $\langle d_1, \ldots, d_n \rangle$ dans $\mathcal{I}$, alors il est vrai (faux) du n-tuple $\langle h_1, \ldots, h_n \rangle$ qui lui correspond dans $\mathcal{I}^*$.

En partant d'une interprétation quelconque, on peut toujours construire une interprétation de Herbrand $\mathcal{I}_H^*$ correspondante, même si les cardinaux de $\mathcal{D}$ et de $H_G$, notés respectivement $\#\mathcal{D}$ et $\#H_G$, diffèrent. Soit la formule $G$ :

(8.12) $\quad G \ : \ \exists x \exists y (x > y),$

et un domaine $\mathcal{D} = \{1, 2, 3\}$. Ici, $H_G = \{a, b\}$. Nous sommes donc dans le cas où $\#H_G < \#\mathcal{D}$. La correspondance constituée par la condition (1) pourra être celle de la figure 8.1.

| $H_G$ | | $\mathcal{D}$ |
|---|---|---|
| $a$ | $\rightarrow$ | 1 |
| $b$ | $\rightarrow$ | 2 |
| | | 3 |

**Figure 8.1** *Cas où* $\#H_G < \#\mathcal{D}$

Si $\|P(d_1, \ldots, d_n)\|_{\mathcal{I}} = \mathbf{V}$ (respectivement $\mathbf{F}$), alors pour n'importe quelle interprétation de Herbrand correspondante $\mathcal{I}_H^*$, $\|P(h_1, \ldots, h_n)\|_{\mathcal{I}_H^*} = \mathbf{V}$ (respectivement $\mathbf{F}$).

Prenons comme prédicat la fonction propositionnelle $x > y$. Les $n$ couples ordonnés qui le vérifient sont soulignés. Ceux qui le falsifient ne sont pas soulignés (cf. figure 8.2).

$\mathcal{I}$

| | | |
|---|---|---|
| 1,1 | 1,2 | 1,3 |
| 2,1 | 2,2 | 2,3 |
| 3,1 | 3,2 | 3,3 |

$\mathcal{I}_H^*$

| | |
|---|---|
| $a,a$ | $a,b$ |
| $b,a$ | $b,b$ |

**Figure 8.2** *Correspondance entre domaines*

On voit que la clause (2) de correspondance est satisfaite : tous les n-tuples de $h_i$, c'est-à-dire $\langle 1,1\rangle, \langle 1,2\rangle, \langle 2,1\rangle, \langle 2,2\rangle$, se comportent vis à vis du prédicat $x > y$ (vérifient ou falsifient celui-ci) comme les n-tuples de $d_i$ auxquels ils correspondent, c'est-à-dire $\langle a,a\rangle, \langle a,b\rangle, \langle b,a\rangle, \langle b,b\rangle$. Il y a donc *correspondance* entre $\mathcal{I}^*_H$ et $\mathcal{I}$, bien que les vérificateurs ou falsificateurs de $x > y$ dans $\mathcal{I}^*_H$ soient moins nombreux que ceux de $\mathcal{I}$ auxquels ils correspondent. La correspondance est asymétrique. Ce n'est pas une bijection; $\mathcal{I}^*_H$ s'aligne sur $\mathcal{I}$ dans les assignations de valeurs de vérité *au prorata* des constantes de Herbrand disponibles.

On peut aussi avoir $\#H_G > \#\mathcal{D}$. Ce sera le cas si $G$ est la formule

(8.13) $\quad \exists x \exists y \exists z \exists u (y > x \wedge z > u),$

et si on se donne $\mathcal{D} = \{1,2,3\}$. En effet, ici $H_G = \{a,b,c,d\}$ puisqu'il y a quatre quantificateurs existentiels en tête et aucun foncteur. La correspondance constituée par (1) pourra être celle de la figure 8.3.

| $H_G$ | | $\mathcal{D}$ |
|---|---|---|
| $a$ | $\rightarrow$ | 1 |
| $b$ | $\rightarrow$ | 2 |
| $c$, $d$ | $\rightarrow$ | 3 |

**Figure 8.3** *Cas où* $\#H_G > \#\mathcal{D}$

Prenons à nouveau comme prédicat $x > y$ : les vérificateurs de ce prédicat sont au nombre de trois dans $\mathcal{I}$ et de cinq dans $\mathcal{I}^*_H$, mais, à nouveau, tous les vérificateurs (falsificateurs) de $\mathcal{I}^*_H$ qui correspondent à des vérificateurs (falsificateurs) de $\mathcal{I}$ alignent leur comportement sur ceux de $\mathcal{I}^*_H$ et, en ce sens, leur correspondent, bien qu'ils soient plus nombreux.

| $\mathcal{I}$ | | | |
|---|---|---|---|
| | $1,1$ | $1,2$ | $1,3$ |
| | $\underline{2,1}$ | $2,2$ | $2,3$ |
| | $\underline{3,1}$ | $\underline{3,2}$ | $3,3$ |

| $\mathcal{I}^*_H$ | | | | |
|---|---|---|---|---|
| | $a,a$ | $a,b$ | $a,c$ | $a,d$ |
| | $\underline{b,a}$ | $b,b$ | $b,c$ | $b,d$ |
| | $\underline{c,a}$ | $\underline{c,b}$ | $c,c$ | $c,d$ |
| | $\underline{d,a}$ | $\underline{d,b}$ | $\underline{d,c}$ | $d,d$ |

**Figure 8.4** *Correspondance entre domaines*

*Remarque.* Il n'y a qu'un seul *domaine de Herbrand* par formule ou ensemble de formules, mais il peut exister plusieurs *interprétations de Herbrand correspondantes* pour une même formule et un même domaine. Dans l'exemple précédent, on aurait pu croiser les deux premières flèches.

### 8.3.5. *Inconsistance sémantique dans le domaine de Herbrand*

**Lemme 8.1.** Si une interprétation $\mathcal{I}$ dans un domaine $\mathcal{D}$ satisfait un ensemble $S$ de clauses, alors n'importe quelle interprétation de Herbrand correspondante $\mathcal{I}_H^*$ le fait aussi dans $H$.

*Démonstration*[5]. Puisque $\mathcal{I}$ satisfait $S$, toutes les clauses de $S$ sont évaluées à **V** par $\mathcal{I}$. Dès lors, par construction de $\mathcal{I}_H^*$, il est évident que toutes les clauses de $S$ seront évaluées à **V** par $\mathcal{I}_H^*$, donc $\mathcal{I}_H^*$ satisfait $S$.

**Lemme 8.2.** Un ensemble $S$ de clauses est sémantiquement inconsistant (c'est-à-dire faux dans toutes les interprétations $\mathcal{I}$) si et seulement si la conjonction de ses clauses est fausse dans toutes les interprétations $\mathcal{I}_H$ des $S$ (interprétations de Herbrand).

Ce lemme est crucial. Il nous permet de nous limiter aux interprétations de Herbrand dans nos preuves d'inconsistance.

*Démonstration.*

1. La direction ($\Rightarrow$) est évidente : ce qui est faux dans *toutes* les interprétations l'est, à fortiori, dans quelques-unes.
2. Pour la direction ($\Leftarrow$) de droite à gauche, on fait une démonstration par l'absurde. Supposons que toutes les interprétations de Herbrand $\mathcal{I}_H$ falsifient la conjonction des clauses de $S$, mais que celle-ci soit néanmoins consistante. Si elle est consistante, il existe au moins un modèle $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$ qui rende vrai la conjonction en question. Mais alors, en vertu du lemme 8.1, il existe une interprétation de Herbrand correspondante $\mathcal{I}_H^*$ qui le fait aussi, ce qui est contraire à l'hypothèse.

### 8.3.6. *La notion de clause fondamentale*

Toute forme de Skolem est logiquement équivalente à une conjonction finie de clauses qui, sauf dans les cas dégénérés, contiennent des variables libres (cf. section 6.4.3). Quand l'univers de Herbrand de la forme de Skolem considérée est infini — ce qui est généralement le cas — il y a une infinité d'interprétations de Herbrand possibles de cette suite finie de clauses. En effet, il y a autant d'interprétations de Herbrand qu'il y a de manières possibles d'assigner comme valeurs aux variables libres des individus puisés dans l'univers de Herbrand de la formule, univers que nous avons supposé infini.

Comme les membres de l'univers de Herbrand de la forme de Skolem sont des symboles et, par surcroît, qu'ils appartiennent à la catégorie des constantes individuelles (celles-ci sont simples (par exemple $a_1, a_2, \ldots$) ou complexes (par

[4] empruntée à P. Delmal

exemple $f(a), f(f(a)) \ldots)$ ), nous pouvons *substituer* ces symboles aux variables individuelles, ce que nous n'aurions pas pu faire (cf. 1.5.2 )si nous avions assigné à ces variables des valeurs prises dans le domaine des individus physiques ou des nombres.

**Définition 8.5 : clause fondamentale.** Une clause dans laquelle on a remplacé toutes les variables libres par des symboles puisés dans l'univers de Herbrand est appelée *clause fondamentale*.

Tantôt nous disions que la suite finie des clauses à laquelle donne lieu l'élimination des quantificateurs universels dans une forme de Skolem admet généralement une *infinité d'interprétations*. Maintenant que nous disposons du concept de clause fondamentale, nous pouvons exprimer ce fait en termes syntaxiques et dire que la suite finie des clauses correspondant à une forme de Skolem admet généralement une *infinité d'instances fondamentales* (on appelle ainsi les clauses fondamentales obtenues en instanciant les variables libres des clauses initiales à l'aide des symboles formant le domaine de Herbrand de la forme de Skolem).

### 8.3.7. *Le théorème proprement dit*

Nous sommes maintenant en mesure de donner une première formulation du théorème de Herbrand.

**Théorème 8.3 : théorème de Herbrand.** Un ensemble $S$ de clauses fondamentales est sémantiquement inconsistant si et seulement s'il existe un ensemble *fini* $S'$ de clauses fondamentales de $S$ *sémantiquement* inconsistant.

Les clauses fondamentales sont des expressions du calcul des propositions. Comme le calcul des propositions est complet, tout ensemble fini de formules du calcul des propositions qui est sémantiquement inconsistant est également syntaxiquement inconsistant. On peut dès lors reformuler le théorème de Herbrand en termes syntaxiques :

(8.14) Un ensemble infini de clauses fondamentales est syntaxiquement inconsistant si et seulement s'il existe un ensemble fini d'instances de telles clauses qui est *syntaxiquement* inconsistant.

Exprimé sous cette forme, le théorème de Herbrand se présente comme un *théorème de compacité syntaxique* (cf. 9.9). Mais il y a plus. Herbrand a montré qu'il est possible de générer mécaniquement cet ensemble fini. Son théorème acquiert de ce fait un caractère *constructif* fort important.

La démonstration du théorème de Herbrand que nous présenterons fait usage de la notion d'*arbre sémantique* et du lemme de König.

### 8.3.8. *Les arbres sémantiques*

**Définition 8.6 : arbres infinis, arbres finitaires, arbres finis.** On distingue les *arbres finis* (figure 8.5), qui n'ont qu'un nombre fini de noeuds ($N_i$), des *arbres infinis* qui en possèdent un nombre infini. Les *arbres syntaxiques* sont des arbres finis. La classe des arbres infinis se subdivise à nouveau en deux sous-ensembles : les arbres *infinis non finitaires* (figure 8.6) et les arbres *infinis finitaires* (figure 8.7). Un arbre infini finitaire est un arbre dont tous les noeuds n'ont qu'un nombre fini de successeurs directs. L'arbre infini *binaire* (figure 8.8) souvent utilisé en informatique est finitaire : chacun de ses noeuds a exactement deux descendants.

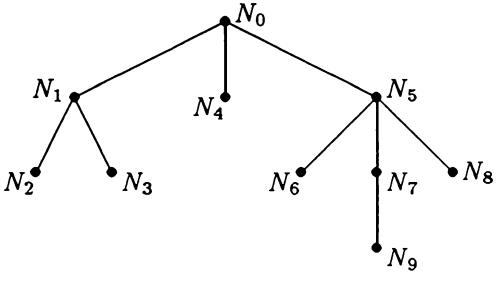



**Figure 8.5** *Arbre fini*

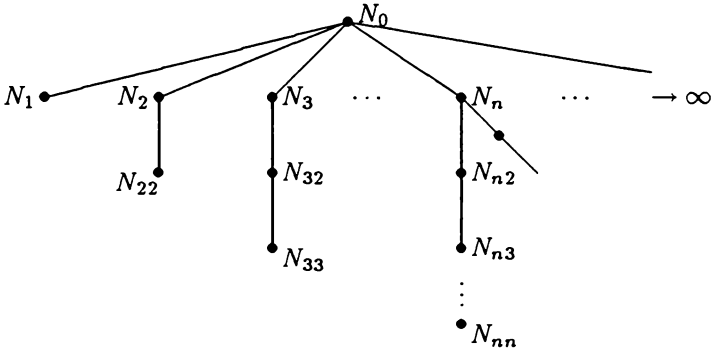



**Figure 8.6** *Arbre infini non finitaire : un noeud au moins a une infinité de successeurs.*

Pour représenter toutes les interprétations possible d'une formule du calcul des propositions comportant $n$ atomes, on peut recourir à un tableau matriciel

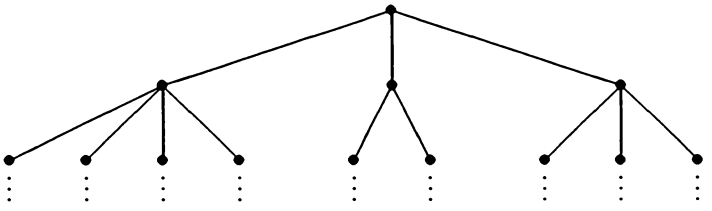


**Figure 8.7** *Arbre infini finitaire*

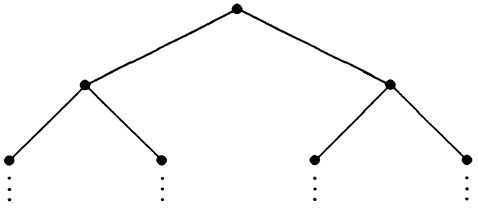


**Figure 8.8** *Arbre binaire (infini)*

comportant $2^n$ entrées (cf. section 3.1.4). On peut également faire usage d'un arbre binaire infini. Au lieu d'écrire

(8.15) $\quad ||p|| = \mathbf{V}$ et $||p|| = \mathbf{F}$,

nous écrirons

(8.16) $\quad p$ et $\overline{p}$.

Chaque branche de l'arbre de la figure 8.9 représente une interprétation possible d'une formule. On notera que l'on peut représenter toutes les interprétations d'une suite infinie de formules à l'aide d'un arbre infini finitaire. On a ici l'équivalent d'un tableau matriciel infini. Le nombre de branches est fonction du nombre de propositions atomiques $(p, q, r, \ldots)$ distinctes comme l'était le nombre de rangées dans un tableau matriciel. On aura dans ce cas une infinité de branches.

**Définition 8.7 : arcs.** Appelons *arcs* les chemins entre deux noeuds entre lesquels il n'y a pas de noeuds. Les arcs correspondent à des symboles de conjonctions.

On voit tout de suite qu'une branche de l'arbre falsifie une clause $C_i$ si la conjonction de littéraux qui correspond à un segment de la branche est la

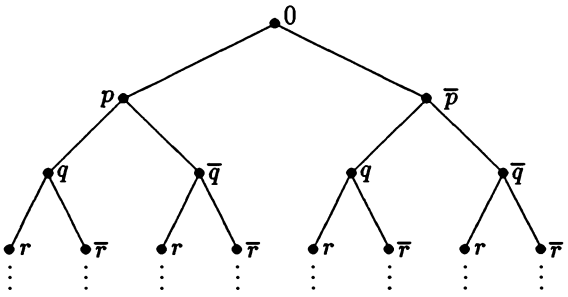



**Figure 8.9** *Arbre d'interprétation*

négation de $C_i$. Par exemple, le segment de branche formé par la conjonction

(8.17) $\overline{p} \wedge q \wedge \overline{r}$,

qui, par abus de langage, est mise pour

(8.18) $\|p\|_{\mathcal{I}} = \mathbf{F}, \|q\|_{\mathcal{I}} = \mathbf{V}, \|r\|_{\mathcal{I}} = \mathbf{F}$,

falsifie la clause

(8.19) $(p \vee \overline{q} \vee r)$.

**Définition 8.8 : base de Herbrand.** On appelle *base de Herbrand* ou *ensemble atomique* pour un ensemble de clauses $\{C_1, \ldots, C_n\}$, l'ensemble $\alpha$ des propositions atomiques figurant dans ces clauses. Il comprend les propositions atomiques fermées ainsi que celles qui étaient initialement ouvertes, mais dans lesquelles on a substitué aux variables libres des constantes puisées dans le domaine de Herbrand $H$ de ces clauses.

**Définition 8.9 : arbre sémantique.** Les *arbres sémantiques* sont des arbres *finitaires* (chaque noeud n'a qu'un nombre fini de successeurs immédiats $Q_1, \ldots, Q_n$) satisfaisant les conditions :

1. La disjonction des étiquettes $Q_i$ des successeurs immédiats d'un même noeud, c'est-à-dire une expression de la forme $Q_1 \vee \cdots \vee Q_n$, est valide.
2. Pour tout noeud $N$, l'union des ensembles de littéraux attachés au segment de branche compris entre la racine et $N$ ne contient aucune paire formée d'un atome et de sa négation.

Sachant que la fourche représente graphiquement la disjonction, et que la succession sur une même branche représente graphiquement la conjonction, on peut reformuler ainsi les conditions ci-dessus :

1. *Lues en parallèle*, les étiquettes des noeuds issus d'un même noeud forment une tautologie.

2. *Lues en série*, les étiquettes des noeuds localisés sur une même branche sont exemptes de contradictions.

**Définition 8.10 : arbre complet.** Un arbre sémantique est dit *complet* s'il satisfait, en plus des conditions exigées de l'arbre sémantique, la condition :

3. Chaque branche, c'est-à-dire chaque séquence d'arcs allant de la racine à un noeud extrémité, contient pour tout atome de la base de Herbrand, soit cet atome soit sa négation.

**Définition 8.11 : noeud-échec.** Un noeud $N$ est appelé *noeud-échec* si la portion de la branche depuis la racine jusqu'à ce noeud falsifie une clause fondamentale instanciée et si aucun segment plus court ne le faisait.

*Remarque.* Cette notion de noeud-échec diffère de celle qui est utilisée dans la méthode des tableaux sémantiques (4.1.16)(6.1). Il faut distinguer les *arbres sémantiques*, dont il est question ici, des *tableaux sémantiques* étudiés précédemment.

**Définition 8.12 : arbre fermé.** Un arbre sémantique complet est *fermé* si et seulement si chacune de ses branches se termine sur un noeud-échec.

*Convention*

Nous adopterons la convention suivante : dès qu'un noeud-échec a été atteint, nous élaguons l'arbre en coupant le reste de la branche.

### 8.3.9. *Construction de l'arbre sémantique complet*

Soit la version implicative du raisonnement introduit à la section 5.3.10. Pour établir la validité de ce conditionnel, nous allons montrer l'inconsistance de sa négation à l'aide d'un arbre sémantique. Nous commençons donc par préfixer la formule à l'aide d'une négation.

(8.20) $\neg((\forall x \forall y \forall z((Fxy \wedge Fyz) \supset Gxz) \wedge \forall x \exists y Fxy) \supset \forall x \exists y Gxy)$.

Mise en forme clausale, la formule donne naissance à un ensemble de trois clauses :

(8.21) $\{\ \overline{F}xy \vee \overline{F}yz \vee Gx,\ Fg(x')x',\ \overline{G}y'a\ \}$.

Le domaine de Herbrand $H$ de cet ensemble de clauses est l'ensemble

(8.22) $\{a, g(a), g(g(a)) \ldots\}$.

Comme $H$ est infini, les trois clauses donnent naissance à une infinité d'instances. On peut se servir d'un arbre sémantique pour énumérer ces instances. On substitue successivement aux variables libres $a, g(a), g(g(a)) \ldots$ pour transformer les *clauses* en *clauses fondamentales*. Convenons — ce qui n'est pas

indispensable, mais commode — d'utiliser un arbre binaire comme arbre sémantique, et élaguons-le dès qu'un segment est la négation d'une clause fondamentale, c'est-à-dire dès qu'il est la négation d'une des clauses de l'ensemble initial, instanciée dans le domaine de Herbrand de cet ensemble de clauses. Pour ce qui est de la formule 8.20, nous obtenons l'arbre de la figure 8.10 en quatre étapes que nous allons décrire.

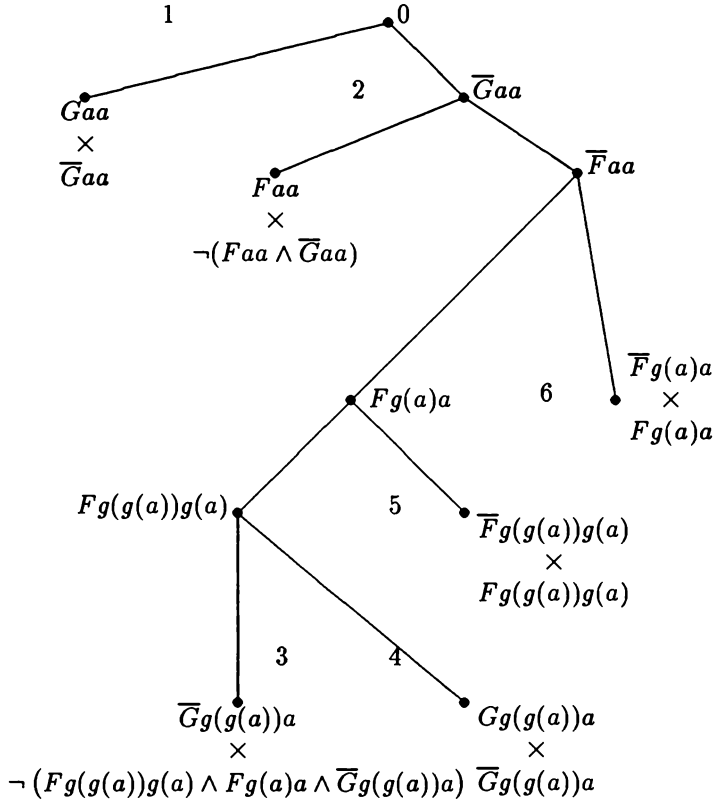



**Figure 8.10** *Arbre sémantique*

Il s'agit d'un arbre fini fermé qui montre toutes les interprétations possibles des trois clauses, c'est-à-dire toutes les instantiations possibles de celles-ci dans le domaine de Herbrand de ces clauses.

Le théorème de Herbrand établit que lorsque toutes les interprétations falsifient les clauses d'une formule mise en forme clausale, l'arbre sémantique binaire associé à $S$ devient, après élagage, un arbre fini fermé.

*Mode de construction d'un arbre sémantique*

On commence par instancier les clauses fondamentales à l'aide de la constante $a$ située au premier niveau, c'est-à-dire $a$. On obtient

$$\underbrace{\overline{F}aa \vee \overline{F}aa \vee Gaa}_{\text{I}}, \underbrace{Fg(a)a}_{\text{II}}, \underbrace{\overline{G}aa}_{\text{III}}$$

On génère l'arbre binaire (figure 8.11).

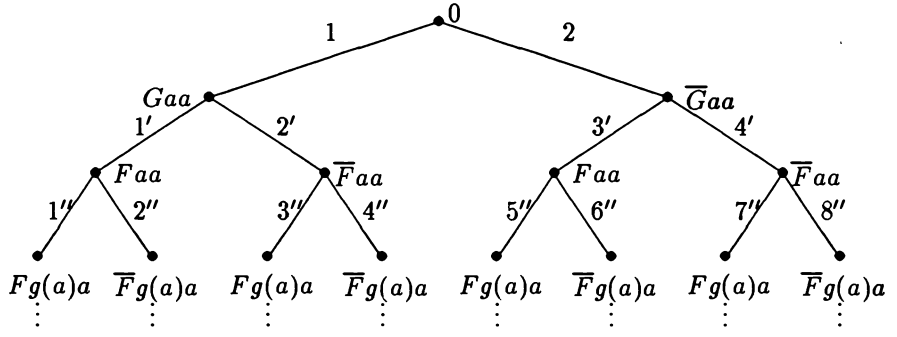



**Figure 8.11** *Arbre sémantique* $H^0$

On élague dès qu'on a un noeud-échec. La clause III permet d'élaguer la branche 1. La clause I, par factorisation (idempotence), se réduit à $\overline{F}aa \vee Gaa$ qui est équivalente à $\neg(Faa \wedge \overline{G}aa)$ par les lois de Morgan. On obtient donc un noeud-échec sur la troisième ligne. La clause II génère un noeud sur la dernière ligne. A ce stade, c'est-à-dire après les trois élagages, nous avons l'arbre de la figure 8.12.

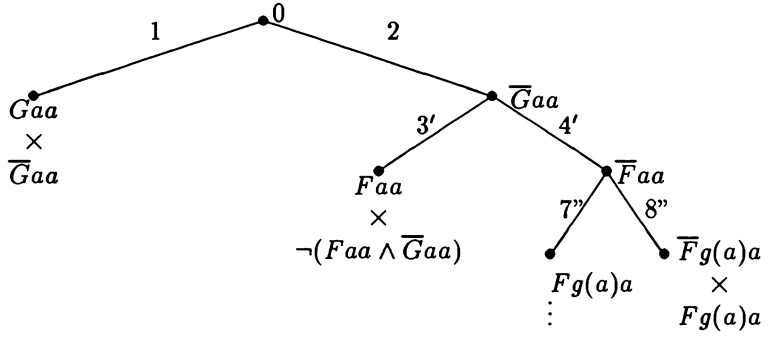



**Figure 8.12** *Arbre sémantique élagué*

On prolonge ensuite la branche ouverte 7″ en puisant dans le niveau suivant de $H$, à savoir dans $H_1 = H_0 \cup \{g(g(a)a)\}$, on obtient en réinstanciant la clause II à l'aide de cette nouvelle constante un nouveau noeud-échec à l'extrémité de 10‴.

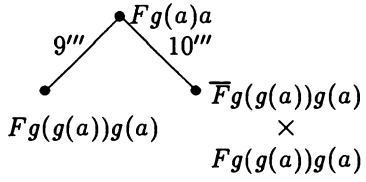



**Figure 8.13**

On prolonge la dernière branche qui reste ouverte (9‴), en ajoutant à son extrémité la bifurcation entre les deux littéraux complémentaires $\overline{G}g(g(a))a$ et $Gg(g(a))a$ et l'on obtient les deux branches de la figure 8.14.

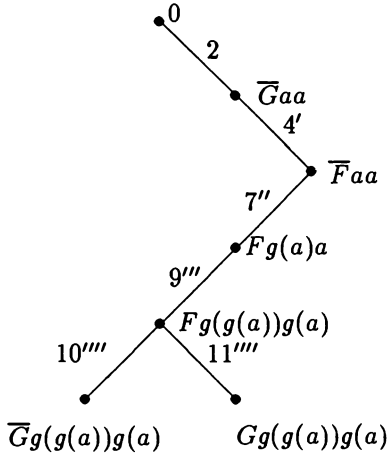



**Figure 8.14** *Arbre sémantique prolongé*

La branche de droite (11⁗) se ferme par la substitution dans III de $g(g(a))$ a $y'$. La branche de gauche (10⁗) se ferme à son tour par la substitution dans I de $g(g(a))$ à $x$, de $g(a)$ à $y$ et de $a$ à $z$. En effet, III devient

$$\overline{F}g(g(a))g(a) \vee \overline{F}g(a)a \vee Gg(g(a))a$$

qui est équivalent, par les lois de Morgan, à

$$\neg(Fg(g(a))g(a) \wedge Fg(a)a \wedge \overline{G}g(g(a))g(a)).$$

Or cette formule est la négation des littéraux associés aux trois derniers noeuds de la branche de gauche, c'est-à-dire aux noeuds reliés par les arcs $10''''$ et $9'''$.

### 8.3.10. *Les quatre étapes de la démonstration de Herbrand*

La démonstration du théorème de Herbrand tient tout entière en quatre propositions :

(8.23) Si $\tau$ est finitaire infini, alors il existe une branche $B$ de $\tau$ infinie.

(8.24) Il existe une branche $B$ infinie si et seulement s'il existe une instantiation qui ne falsifie pas une clause.

(8.25) Toutes les instantiations falsifient une clause.

(8.26) Donc l'arbre $\tau$ finitaire est fini.

La proposition 8.23 n'est autre que le lemme de König.

### 8.3.11. *Lemme de König*

**Lemme 8.4 : lemme de König (1926).** Chaque arbre finitaire infini a au moins une branche infinie.

*Démonstration.* Appelons *immortel* un noeud qui a une infinité de successeurs (immédiats ou non) et *mortel* un noeud qui n'a qu'un nombre fini de successeurs. Comme l'arbre considéré est infini, il a *au moins* un noeud immortel : le noeud origine $N_0$. Comme, par surcroît, il est finitaire, les successeurs immédiats de $N_0$ comptent au moins un noeud immortel $N_1$. En effet, si ce n'était pas le cas, c'est-à-dire si tous leurs descendants étaient en nombre fini, l'arbre serait fini puisque les successeurs immédiats de $N_0$ sont en nombre fini. Mais ce raisonnement vaut également pour $N_1$, le successeur immortel de $N_0$. Ses successeurs aussi sont en nombre fini et il doit donc exister *au moins* un noeud immortel parmi eux, appelons le $N_2$. La suite $\langle N_0, N_1, N_2, N_3 \ldots \rangle$ constitue une branche infinie.

### 8.3.12. *Démonstration détaillée du théorème de Herbrand*

**Définition 8.13.** Un arbre $\tau'$ est dit *correspondre* à l'arbre $\tau$ s'il est obtenu à partir de $\tau$ par *élagage*.

**Théorème 8.5 : théorème de Herbrand.** Un ensemble $S$ de clauses est sémantiquement inconsistant si et seulement si à chaque arbre *sémantique complet* $\tau$ de $S$ correspond un arbre *sémantique*[6] *fermé fini* $\tau'$.

*Démonstration.* Démontrons d'abord l'implication de gauche à droite.

Supposons $S$ sémantiquement inconsistant et $\tau$ un arbre sémantique complet de $S$. Pour chaque branche $B$ de $\tau$, soit $\mathcal{I}_B$ l'ensemble de tous les littéraux attachés aux segments de la branche $B$. Tous les $\mathcal{I}_B$ sont des interprétations de $S$, et, comme $S$ est sémantiquement inconsistant, tous les $\mathcal{I}_B$ sont la négation d'une clause fondamentale $C'$ instanciant une clause $C$ de $S$. Comme toutes les clauses de l'ensemble éventuellement infini $S$ sont finies, il doit exister un noeud-échec $N_B$ sur chaque branche $B$, donc on obtient un arbre $\tau'$ fermé. Par définition, dans un arbre sémantique, tout noeud n'a qu'un nombre fini de successeurs immédiats. Par surcroît, $\tau'$ est fini car s'il ne l'était pas, il aurait — en tant que finitaire — une branche infinie, en vertu du lemme de König. Dans cette hypothèse, il y aurait une branche ne contenant aucun noeud-échec, donc l'arbre ne serait pas fermé, ce qui est impossible en vertu de la première partie.

La réciproque est immédiate. Si pour tout arbre sémantique complet de $S$, il existe un arbre sémantique fini fermé qui lui correspond, alors toute branche contient un noeud-échec et toute interprétation $\mathcal{I}_B$ rend $S$ faux, donc $S$ est sémantiquement inconsistant.

### 8.3.13. *Le constructivisme de Herbrand*

Pour construire mécaniquement l'arbre sémantique d'une formule en forme clausale, on doit étiqueter les noeuds d'un arbre binaire avec les atomes de la base de Herbrand de l'ensemble des clauses fondamentales de la formule, ces clauses étant obtenues en substituant aux variables libres les symboles figurant dans le domaine. (On se souvient que le domaine de Herbrand d'une formule est constitué de symboles, c'est-à-dire d'entités linguistiques.) Chemin faisant, on élague l'arbre chaque fois qu'une branche a crû assez pour falsifier une clause.

J. van Heijenoort compare la construction d'une interprétation à partir d'un arbre à la notion de déploiement chez Brouwer :

> Pour chaque interprétation qui se trouve à l'un des noeuds de l'arbre, un symbole de prédicat $P$ a la valeur V pour un certain nombre fini de termes, et la valeur F pour un nombre fini d'autres termes. Lorsqu'on s'avance le long d'une branche infinie de l'arbre, ces deux ensembles finis de termes croissent (ou du moins ne décroissent pas) et peuvent être considérés comme des approximations finies de plus en plus fines des deux

[6] Rappelons que tout arbre sémantique est finitaire

ensembles associés, l'un à $P$, l'autre à $\neg P$, dont l'union forme le domaine de l'interprétation-limite. Nous rejoignons ici la notion de déploiement telle que Brouwer l'avait introduite pour fonder l'analyse intuitionniste et qu'il considérait comme devant se substituer à la notion d'ensemble, trop suspecte (van Heijenoort, 1982, pp. 76–77).

### 8.3.14. *De Herbrand à Henkin*

J. van Heijenoort a proposé une classification qui éclaire les rapports entre le théorème de Herbrand et le théorème de complétude. Il distingue trois méthodes : la méthode axiomatique ($A$), la méthode de Herbrand ($H$) et la méthode ensembliste ($E$).

> Appelons *méthode axiomatique* celle qui consiste à développer la logique, à la manière de Frege et de Russell, à partir d'axiomes et de règles formelles. Appelons *méthode ensembliste* celle qui fait reposer la logique sur des considérations appartenant à la théorie des ensembles. Appelons enfin *méthode de Herbrand*, celle qui aborde les problèmes logiques par l'étude de la conjonction ou disjonction de Herbrand des formules (van Heijenoort, 1968, pp. 6–7).

Soit $G$ une formule quelconque du calcul des prédicats. Considérons les trois énoncés métalogiques :

(8.27) $A_1$ : $G$ est démontrable par la méthode axiomatique,

(8.28) $E_1$ : $G$ est valide,

(8.29) $H_1$ : il existe un nombre $r$ tel que la $r$-ième disjonction de $G$ est valide en logique des propositions.

Le théorème fondamental de Herbrand établit que $A_1$ et $H_1$ s'impliquent mutuellement. Herbrand savait, par ailleurs, que $E_1$ implique $H_1$, et van Heijenoort avance des considérations justifiant l'implication inverse de $E_1$ par $H_1$. Si on met bout à bout

(8.30) $A_1$ est équivalente à $H_1$,

(8.31) $H_1$ est équivalente à $E_1$,

on obtient tout de suite, par transitivité de l'équivalence, le théorème de complétude

(8.32) $A_1$ est équivalente à $E_1$.

Herbrand ne s'est pas interrogé sur les rapports entre $A$ et $E$. Son finitisme lui faisait considérer comme inacceptable la notion ensembliste de validité :

> [Herbrand] considère que ces notions doivent être remplacées par celles qu'il a introduites ($H_1$), et qui constituent la forme précise et seule

légitime d'énoncer $E_1$ [...], un peu à la manière dont la définition de fonction continue à l'aide de $\epsilon$ et $\delta$ doit se substituer à l'intuition vague d'une courbe sans interruption (van Heijenoort, 1968, p. 8).

Dans le mémoire intitulé *Axiomatische Untersuchung des Aussagenkalküls* der P.M. (*Principia Mathematica*) (1926), dû à P. Bernays, on trouve une démonstration de la complétude du calcul propositionnel (toute formule valide de ce calcul s'obtient à partir des axiomes donnés dans les *Principia*). En 1930, Gödel étend ce résultat au "calcul fonctionnel restreint", c'est-à-dire au calcul des prédicats du premier ordre sans égalité. On en trouve une version claire et simplifiée dans Novikov (tr. fr. 1964). En 1949, Henkin a publié une autre démonstration de complétude du calcul des prédicats du premier ordre, qui constitue un progrès par rapport à celle de Gödel sur deux points :

1. Elle peut être généralisée à des systèmes qui contiennent une infinité non dénombrable de symboles primitifs. Comme le note Henkin lui-même, cette généralisation est sans intérêt pour quiconque voit dans la logique un instrument d'analyse du langage. En revanche, elle est intéressante pour le mathématicien, en raison des applications qu'elle a en algèbre abstraite.
2. Elle peut être étendue à des systèmes de logique supérieurs au premier ordre (logique du deuxième ordre, théorie des types) à condition de restreindre le concept de validité.

   Alors qu'en logique du premier ordre, Gödel et Henkin ont prouvé que toute formule valide (vraie dans tous les modèles) est démontrable, en logique du deuxième ordre, Henkin a prouvé le théorème suivant :

   > Si on entend par modèle général un modèle où les domaines d'objets de type supérieur ne contiennent pas *tous* les sous-ensembles du domaine de base (ou respectivement tous les sous-ensembles des produits cartésiens de ce domaine), et si on définit comme formule valide du second ordre une formule valide dans tout domaine général, alors la logique du second ordre est complète relativement aux modèles généraux (Largeault 1972, p. 199).

Pour une présentation claire et simplifiée du théorème de complétude de Henkin pour le calcul des prédicats du premier ordre formalisé dans système axiomatique à la Hilbert, voir Rivenc 1989. Il existe également une preuve de complétude pour la méthode des tableaux sémantiques (Beth, Hintikka, Smullyan); voir également Rivenc.

Indépendamment, Birkoff (1935) a démontré la complétude d'un *fragment* de la logique des prédicats du premier ordre, à savoir le calcul équationnel, dans lequel les règles d'inférence sont limitées à des règles qui ne font référence qu'à des équations dans leurs prémisses et leurs conclusions. L'intérêt de ce fragment pour l'informaticien est le suivant :

> Le théorème de correction (*soundness*) et complétude pour le calcul équationnel révèle un fait intéressant et utile : les règles plus puissantes du calcul des prédicats du premier ordre avec égalité[7], appliquées à une théorie équationnelle, ne génèrent pas de nouvelles équations qui ne pouvaient pas être dérivées à l'aide du raisonnement purement équationnel. Comme le calcul équationnel est beaucoup plus simple [que la logique du premier ordre avec égalité], il présente un intérêt et un attrait en tant que calcul à automatiser au moyen des techniques de démonstration à la machine (Abramsky et al., à paraître).

La version du théorème de complétude que nous présentons s'applique au système de preuve par séquents. Rappelons que le théorème de complétude faible se formule ainsi :

$$\text{Si } \models \Phi \text{ alors } \vdash \Phi.$$

Il est la réciproque du théorème de correction (*soundness*) appelé parfois adéquation :

$$\text{Si } \vdash \Phi \text{ alors } \models \Phi.$$

## 8.4. Complétude pour les séquents du premier ordre

### 8.4.1. *Préliminaire*

Soient $E_1, \ldots, E_k, F_1, \ldots, F_l$ des formules du calcul des prédicats ne contenant pas de symboles fonctionnels[8] et ne contenant pas de variable liée qui figure libre dans l'une d'elles[9], soit le séquent

(8.33) $\quad E_1, \ldots, E_k \Longrightarrow F_1, \ldots, F_l,$

qui est la racine d'un arbre que l'on a fait croître vers le haut en appliquant les quatorze règles (dix pour les connecteurs : 4.1.2.3, quatre pour les quantificateurs : 6.2.2) et le plan de travail. Trois cas sont possibles.

---

[7] On obtient le calcul des prédicats avec identité en ajoutant aux axiomes du calcul des prédicats respectivement l'axiome et le schéma d'axiome suivants :

$$\forall x(x = x),$$
$$\forall x \forall y(x = y \supset (\Phi x \supset \Phi y)).$$

*Exercice* : déduire de ces axiomes la symétrie et la transitivité de l'identité ainsi que la propriété exprimée par

$$\forall x \forall y \forall z(((x = z) \wedge (y = z)) \supset x = y)$$

[8] Cette limitation n'est pas vraiment un appauvrissement, dans la mesure où l'on peut éliminer les symboles fonctionnels grâce à l'emploi de prédicats.

[9] Quand une même variable figure à la fois libre et liée, un renommage de la variable liée met fin à cet état de choses.

1. Il existe un arbre fini clos. Dans ce cas, le séquent est démontrable. En effet, si l'arbre se ferme, c'est que chacune de ses branches contient à son extrémité un axiome, et il suffit — en vertu de la réversibilité des règles (cf. section 4.1.4) — de relire l'arbre de preuve de haut en bas pour obtenir une preuve du séquent placé à la racine.
2. Il existe un arbre fini ouvert. Dans ce cas, nous montrerons dans un lemme que le séquent placé à la racine est falsifiable.
3. Il existe un arbre infini ouvert. Dans ce cas, le lemme précité prouve également que le séquent est falsifiable.

Une fois le lemme démontré, nous pourrons écrire

(8.34) Ou le séquent 8.33 est falsifiable, ou il est démontrable,

ce qui est trivialement équivalent à

(8.35) Si le séquent 8.33 n'est pas falsifiable — en d'autres termes, s'il est *valide* — alors il est *démontrable*.

### 8.4.2. *Lemme*

**Lemme 8.6.** Dans un arbre de séquents construit de bas en haut à partir de $E_1, \ldots, E_k \Longrightarrow F_1, \ldots, F_l$, conformément aux quatorze règles (introduction et élimination de $\neg, \wedge, \vee, \supset, \equiv, \exists, \forall$) et au plan décrit, à tout chemin non clos terminé ou infini correspond un contre-exemple à $E_1, \ldots, E_k \Longrightarrow F_1, \ldots, F_l$, de domaine $\mathcal{D} = \{0, \ldots, q\}$ ou $\mathcal{D} = \{0, 1, 2 \ldots\}$ selon que $t_0, \ldots, t_q$ ou $t_0, t_1, t_2 \ldots$ sont activés le long de ce chemin (Kleene, 1967, 1972, p. 305).

**Définition 8.14 : ancêtre et descendant.** L'arbre est orienté comme dans la nature, la racine en bas, mais on peut aussi le voir comme un arbre généalogique. Si on adopte cette deuxième perspective, on peut dire que chaque formule de type $\alpha$ provient d'un *ancêtre immédiat*, et chaque proposition de type $\beta$, de deux *ancêtres immédiats* situés sur la ligne immédiatement au-dessus. On peut dire qu'elle est la *descendante* immédiate de ces ancêtres. Les formules situées au sommet seront dites ancêtres d'elles-mêmes.

*Rappel.* Quand on suit le plan de travail de la section 6.2.2, chaque formule de type $\alpha$, $\beta$ ou $\delta$ intervient comme formule principale (c'est-à-dire comme formule sur laquelle on travaille) exactement une fois sur un chemin.

*Démonstration du lemme 8.6 (adaptée de Kleene 1967, 1972).* On considère un chemin (une branche) *non clos*. Le long de ce chemin se trouve une suite finie ou infinie de séquents. On forme l'union de toutes les formules figurant dans les antécédents de ces séquents, et on l'appelle $\cup\Delta$. On forme l'union de toutes les formules figurant dans les succédents de ces séquents, et on l'appelle $\cup\Lambda$. On va montrer qu'il y a moyen de trouver un modèle $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$ qui rende vrais tous les membres de $\cup\Delta$, et faux tous les membres de $\cup\Lambda$.

Il n'y a pas d'atome commun à $\cup\Delta$ et $\cup\Lambda$. En effet, une fois qu'un atome $C$ est apparu dans un antécédent ou dans un succédent, par suite de l'élimination d'un connecteur ou de l'instantiation d'un quantificateur, il remonte le long de la branche où il est apparu. Soit $\Delta_a$ le premier antécédent dans lequel $C$ est apparu et $\Lambda_b$ le premier succédent dans lequel $C$ est apparu. Si $C$ est à un moment donné à la fois dans $\Delta_c$ et dans $\Lambda_c$, où $c = \max(a,b)$, il y reste et dès ce moment au moins le chemin serait terminé clos, ce qui est contraire à l'hypothèse.

**Première étape :** interprétation des termes, des propositions atomiques et des prédicats.

1. On commence par associer à chaque *terme singulier* (variable ou constante) un individu du domaine :

$$\mathcal{I} \ : \ \{t_0, \ldots, t_q\} \mapsto \{0, \ldots, q\},$$

et si les termes singuliers sont en nombre infini,

$$\mathcal{I} \ : \ \{t_0, t_1 \ldots\} \mapsto \mathbf{N}.$$

2. On donne ensuite une valeur de vérité aux symboles de propositions atomiques ($P$, $Q$, $R\ldots$), en se conformant à la règle suivante :

| | |
|---|---|
| Si $P \in \cup\Delta$, | $\|\|P\|\|_{\mathcal{I}} = \mathbf{V}$, |
| Si $P \in \cup\Lambda$, | $\|\|P\|\|_{\mathcal{I}} = \mathbf{F}$, |
| Si $P \notin \cup\Delta$ et $P \notin \cup\Lambda$, | $\|\|P\|\|_{\mathcal{I}} =$ une valeur arbitraire, choisissons $\mathbf{F}$. |

3. Les prédicats à $n$-places ($F(-)$, $G(-,-)$, $H(-,-,-)\ldots$) qui figurent dans des propositions atomiques (analysées en "sujets" et prédicats), telles que $Fx$, $Gay$, $Hadb\ldots$, seront interprétés à l'aide d'une *fonction logique*[10], appelée aussi *fonction caractéristique* $f$ qui prend comme valeur $\mathbf{V}$ pour les termes $t_1, \ldots, t_q$ ou $t_1, t_2 \ldots$ qui, substitués aux places vides $(-)$, $(-,-)$, $(-,-,-)$ du prédicat, engendrent des propositions atomiques figurant dans $\cup\Delta$, et $\mathbf{F}$ pour ceux qui engendrent des propositions figurant dans $\cup\Lambda$ ou ne figurant ni dans $\cup\Delta$, ni dans $\cup\Lambda$.

**Deuxième étape :** interprétation des propositions moléculaires et des propositions générales.

On va montrer que le modèle $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$ choisi pour rendre vraies toutes les propositions atomiques de $\cup\Delta$ et pour rendre fausses les propositions atomiques de $\cup\Lambda$ rend également vraies les autres propositions. Supposons le contraire. En d'autres termes supposons qu'une ou plusieurs propositions fassent exception à la règle et soient vraies alors qu'elles sont dans $\cup\Delta$ ou fausses alors qu'elles sont dans $\cup\Lambda$. Choisissons parmi ces formules celle qui est la plus simple

[10] cf 2.3.5

(qui contient le moins d'occurrences d'opérateurs possible, mais au moins une occurrence).

Considérons d'abord le cas où l'opérateur principal de cette formule est un connecteur ou un quantificateur existentiel (ou quasi existentiel).

Par hypothèse inductive, on suppose le lemme vérifié pour toutes les formules de complexité $n-1$. En d'autres termes, les formules-ancêtres immédiats de $G$ ont la valeur désirée, c'est-à-dire **V** si elles sont dans $\cup\Delta$ et **F** si elles sont dans $\cup\Lambda$. Mais dans ce cas $G$ doit l'avoir aussi — contrairement à ce que nous avons supposé — étant donné que les règles sont correctes et transmettent la valeur désirée.

Si l'opérateur principal de $G$ est un quantificateur universel (ou quasi universel), la formule $G$ n'a pas une formule-ancêtre ou deux, mais une multiplicité éventuellement infinie de formules-ancêtres. Mais comme ces dernières sont toutes de complexité moins grande, elles ont la valeur désirée par hypothèse inductive. Dès lors, à nouveau, $G$ doit l'avoir aussi — pour la raison donnée plus haut — contrairement à ce que nous avons supposé.

*Exemple.* Si $G$ est une universelle qui appartient à $\cup\Delta$ — ce qui veut dire qu'elle est du type $\forall \Longrightarrow$ —, alors les formules-mères $A(t_0)$, $A(t_1)$, $A(t_2)\ldots$ ont la valeur **V** par hypothèse inductive et, les règles d'inférence étant correctes, $\forall x A$ doit avoir aussi la valeur **V**, contrairement à l'hypothèse.

*Remarque.* Dans ce lemme, on a utilisé la loi de Clavius, qui est la réciproque du principe de la preuve par l'absurde

(8.36) $\quad (\neg p \supset p) \supset p.$

En effet, on a dit : "Si $G$ ne possède pas la valeur désirée, elle la possède tout de même, donc elle la possède."

### 8.4.3. *Exemple*

Considérons un arbre de séquents comportant une branche ouverte, et montrons qu'il est possible d'interpréter les symboles et les formules de manière telle que tous les membres de $\cup\Delta$ aient la valeur **V** et que tous les membres de $\cup\Lambda$ aient la valeur **F**(cf. figure 8.15).

On voit aisément que le modèle $\mathcal{M} = \langle \mathcal{D}, \mathcal{I} \rangle$, avec

$$
\begin{array}{rcll}
\mathcal{D} & = & & \{1,2\} \\
\|p\|_{\mathcal{I}} & = & & \mathbf{V} \\
\|b\|_{\mathcal{I}} & = & & 1 \\
\|c\|_{\mathcal{I}} & = & & 2 \\
\|Q\|_{\mathcal{I}} & = & f : & 1 \mapsto \mathbf{V} \\
 & & & 2 \mapsto \mathbf{F}
\end{array}
\tag{8.37}
$$

falsifie

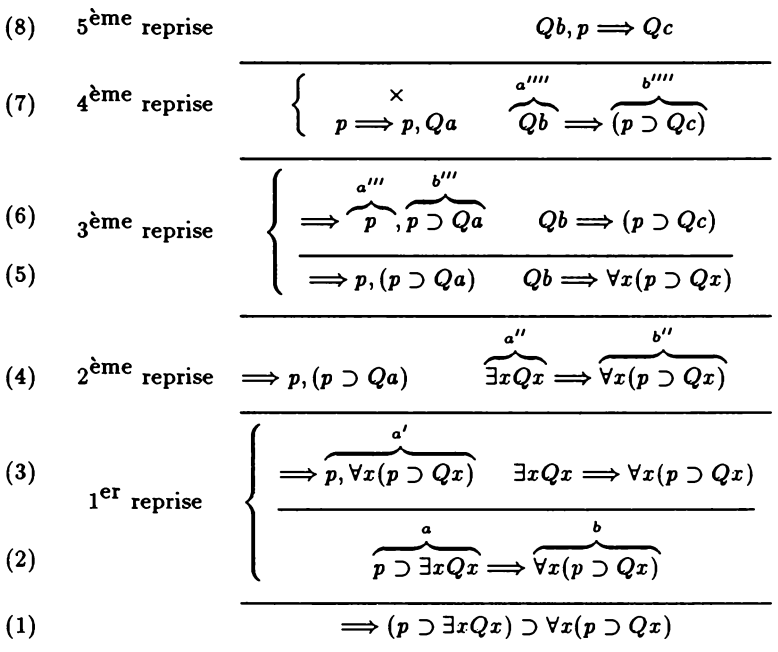



**Figure 8.15** *Arbre ouvert*

(8.38) $(p \supset \exists x Qx) \supset \forall x(p \supset Qx)$.

### 8.4.4. *Construction quasi-mécanique*

Le plan de travail permet soit de démontrer *mécaniquement* la forme si elle est valide, soit de construire un contre-exemple si elle ne l'est pas. Dans le cas où l'arbre est fini ouvert, la construction du contre-exemple est mécanique. Dans le cas où l'arbre est infini ouvert, le plan de travail nous libère d'un effort inventif, mais il ne constitue pas un algorithme : un contre-exemple infini ne peut être construit pas à pas, en un nombre fini d'actions. Le plan de travail est donc quasi-mécanique, plutôt que mécanique au sens strict. Serait-il possible de construire le contre-exemple de manière algorithmique ? Une réponse négative à cette question posée par Hilbert a été donnée en 1936 par le théorème d'*indécidabilité* de Church, que nous allons étudier à présent.

## 8.5. Problème de l'arrêt et problème de correspondance

Il existe des formules

(8.39) $\forall x \exists y \forall z (\overline{R}xx \wedge Rxy \wedge (Ryz \supset Rxz)$

qui n'admettent de contre-exemple que dans un domaine infini. C'est Löwenheim qui a constaté en 1915 l'existence de telles formules. Ces formules sont-elles *réfractaires* à la *détection mécanique* parce qu'elles ne sont vraies que dans un modèle dont le domaine est *infini* ? L'explication est plus subtile :

> Dans le cas particulier de cette formule [$\neg F$ c'est à dire 8.39], le fait qu'il existe un modèle infini est évident (on a par exemple le modèle arithmétique $\mathcal{D} = \mathcal{N}$, $Rxy = x < y$) et on pourrait concevoir une machine programmée de telle sorte qu'elle donne la réponse (négative, si la question est celle de la validité de **F**). Mais la variété des formules de ce type est indéfinie (celle-ci est l'une des plus simples), et un procédé qui réussirait avec l'une échouerait avec d'autres : il n'y a pas de *critère général* qui permettrait de reconnaître dans tous les cas, en examinant son mode de croissance, qu'un arbre ne se révélera jamais fermé, si loin qu'on poursuive sa construction (donc pas de machine à tester la validité qui puisse, dans tous les cas, mettre elle-même un terme à son fonctionnement) (B. Ruyer 1990, p.178)

Cet argument reste valable, même si l'on adopte la méthode de preuve plus perfectionnée de Kleene, c'est-à-dire la méthode des séquents accompagnée du plan de travail. En effet, le plan de travail nous permet de construire un contre-exemple à $\neg F$, c'est-à-dire un modèle vérifiant $F$) dès que nous *savons* que l'arbre des séquents contient un chemin fini non clos ou un chemin infini, mais nous ne disposons de cette information que dans des cas privilégiés comme par exemple pour les formules pour lesquelles la recherche d'un contre-exemple nous conduit à répéter indéfiniment la même succession d'opérations, par exemple instantier $\forall x \exists y Fxy$ avec $a$ puis $y$ avec $b$, puis réinstancier $x$ avec $b$ et $y$ avec $c$ et ainsi de suite.

Pour prouver qu'*en général* nous ne pouvons pas savoir *à l'avance* si une branche ouverte se fermera, ces considérations empiriques ne suffisent pas car elles pourraient, comme le note encore Ruyer, être imputables à la méthode logique particulière que l'on utilise. Le problème posé par Hilbert au début du siècle a été résolu en 1936 par A. Church qui a démontré que le calcul des prédicats n'est que semi-décidable : l'ensemble des schémas valides (et celui des schémas inconsistants) est récursivement énumérable, mais il n'est pas récursif. Son complément, l'ensemble des schémas non valides, n'est pas récursivement énumérable (2.3.5).

La démonstration que nous présenterons et qui est due à Manna peut se résumer ainsi :

On considère trois questions.

(1) Soit $X$ une formule du calcul des prédicats. Trouver un algorithme qui détermine si elle est valide ou non valide.

(2) Soit $Y$ un ensemble de paires de syllabes différentes

$\langle a_1\ b_1\rangle,\ \langle a_2 b_2\rangle, \ldots, \langle a_n\ b_n\rangle$

Trouver un algorithme qui détermine si en prenant dans un certain ordre les $a_i$ et dans le même ordre les $b_i$ avec répétition, il est possible de former deux échantillons du même mot (deux "*tokens*" du même "*type*" au sens de Peirce; cf. 1.5.1).

(3) Soit $Z$ le programme d'un automate abstrait (machine de Post). Trouver un algorithme qui détermine si une machine de Turing $x$ à laquelle on donne comme entrée $y$ s'arrêtera ou si elle continuera indéfiniment son calcul.

On montrera que le problème (1) est au moins aussi difficile que le problème (2) et que le problème (2) est au moins aussi difficile que le problème (3) lequel a été prouvé insoluble.

### 8.5.1. *Problèmes et instances de problèmes*

**Définition 8.15 : problème.** On appelle ici problème une classe infinie dénombrable de questions dichotomiques qui sont les instances d'un même schéma, comme "7 est-il pair ?" est une instance de "$x$ est-il pair ?" (où $x$ a comme domaine les entiers).

En mathématiques, on s'interroge sur la question de savoir s'il existe un algorithme qui permet de résoudre *partiellement* ou *totalement* un *problème* au sens qu'on vient de définir.

### 8.5.2. *Machines de Post*

Les machines de Post[11] appartiennent à la famille des automates abstraits qui permettent de donner un contour précis à la notion intuitive de calculabilité. Comme les fonctions récursives, ces machines sont l'expression formelle de l'idée d'un calcul qui se fait pas à pas. Elles ne sont pas sujettes aux limitations des calculateurs humains ou des ordinateurs, dont la mémoire est finie.

Le problème s'énonce ainsi. On présente un mot formé de $a$ et de $b$ et on désire savoir si

le nombre de $a$ = le nombre de $b$ ou le nombre de $a$ diffère-t-il du nombre de $b$ d'une unité ?

On veut pouvoir répondre à la question *sans* procéder à un comptage. En effet, le recours à un comptage exigerait une mémoire infinie puisqu'il n'y a pas de borne supérieure à la longueur des mots.

[11] La majeure partie de la section 8.5.2 est due à M. P. A. de Marneffe.

On remplace le comptage par une autre procédure : les manipulations permises à une machine de Post se réduisent à celles qui sont décrites dans deux sortes d'instructions :

1) *les instructions-tests :*

effacer $a$ en tête du mot, ce qu'on note *del(a).*

effacer $b$ en tête du mot, ce qu'on note *del(b).*

2) *les instructions-affectations :*

ajouter $a$ à la queue du mot, ce qu'on note *in(x,a).*

ajouter $b$ à la queue du mot, ce qu'on note *in(x,b).*

Il y a deux états terminaux : ACCEPT et REJECT. La machine s'arrête sur ACCEPT si le mot comporte autant de $a$ que de $b$, et en particulier sur le mot "vide". Elle s'arrête sur REJECT si le nombre de $a$ diffère du nombre de $b$ d'une unité, et en particulier si le mot est constitué par une seule lettre ($a$ ou $b$)

```
DO  {B1}  x = ∅      →  ACCEPT
□   {B1}  del(a,x)   →  DO  {B2}  del(a,x)  →  {B4} in(x,a) {B2}
                        OD;
                        IF  {B2}  x = ∅      →  REJECT
                        □   {B2}  del(b,x)   →  skip{B1}
                        FI
□   {B1}  del(b,x)   →  DO  {B3}  del(b,x)  →  {B5} in(x,b) {B3}
                        OD;
                        IF  {B3}  x = ∅      →  REJECT
                        □   {B3}  del(a,x)   →  skip{B1}
                        FI
OD
```

**Figure 8.16** *Machine de Post détectant la (quasi) parité*

Appelons $N_a$ et $N_b$ respectivement le nombre de $a$ et le nombre de $b$ dans *le ruban initial*. Appelons $n_a$ et $n_b$ respectivement le nombre de $a$ et le nombre de $b$ dans *un état quelconque du ruban* pendant le traitement par la machine.

Dans l'état $B_1$, la condition $(n_a - n_b = N_a - N_b)$ est vérifiée. Cette condition est en fait l'invariant de la boucle externe : elle est vérifiée initialement, et le programme doit être construit afin de la garantir à l'issue de toute itération.

- Si, dans l'état $B_1$, la condition $(x = \emptyset)$ est également respectée, on a $(n_a - n_b = 0)$, d'où $(0 = N_a - N_b)$, c'est-à-dire que la machine peut annoncer "ACCEPT".

- Si, dans l'état $B_1$, le ruban est non vide et préfixé par un $a$, la fonction $del(a, x)$ est 'vrai' et a pour effet d'effacer ce caractère $a$. La machine passe à l'état $B_2$ où la condition suivante est vérifiée : $(n_a + 1 - n_b = N_a - N_b)$.
- Dans l'état $B_2$, si le ruban est vide, la condition devient : $(1 = N_a - N_b)$, c'est-à-dire que la machine peut annoncer "REJECT".
- Si, dans l'état $B_2$, le ruban est non vide, il est préfixé par une suite (éventuellement vide) de caractère $a$. La boucle interne associée à l'état $B_2$ a pour effet de recopier cette suite de $a$ en fin de ruban. Cette copie ne modifie pas le nombre de caractères du ruban et la condition $(n_a + 1 - n_b = N_a - N_b)$ reste vérifiée.
- Si dans l'état $B_2$, le ruban est non vide et ne contient pas de caractère $b$, la boucle interne est activée; mais elle ne se terminera pas puisque seule la rencontre d'un caractère $b$ peut falsifier le gardien "$del(a, x)$". Dans ce cas, la condition vraie en $B_2$ devient $(n_a + 1 = N_a - N_b)$, avec $(n_a \geq 1)$; la spécification du problème est respectée : la machine boucle car l'excès du nombre de $a$ sur le nombre de $b$ *dans le ruban initial* est au moins supérieur à l'unité.
- Considérons, dans l'état $B_2$, un ruban non-vide et contenant au moins un caractère $b$. La boucle interne est activée si le premier caractère $b$ du ruban est précédé par une suite non vide de caractères $a$. Cette activation éventuelle de la boucle interne se terminera vu l'existence d'au moins un caractère $b$ dont la rencontre falsifie le gardien. Ce carctère $b$ est effacé (par $del(b, x)$) dans la commande de choix). La condition $(n_a + 1 - n_b = N_a - N_b)$ devient, après effacement du caractère $b$ : $(n_a + 1 - (n_b + 1) = N_a - N_b)$, soit : $(n_a - n_b = N_a - N_b)$, condition identique à la condition qui doit être vraie en $B_1$. L'invariant de la boucle externe est bien vérifié.

L'activité du programme dans l'état $B_1$ avec un ruban préfixé par un caractère $b$ conduit à la même conclusion après analyse similaire. (La condition en $B_3$ est $(n_a - (n_b + 1) = N_a - N_b)$).

La présentation donnée ci-dessus associe explicitement des conditions aux états. Ces conditions montrent que, tout au long du traitement, la machine "connaît" implicitement la valeur de $(N_a - N_b)$.

La machine boucle dès que l'écart entre le nombre de $a$ et le nombre de $b$ dépasse l'unité. Montrons-le en appliquant le programme au mot $x = aaab$. Nous décrirons les étapes du calcul à l'aide de paires de la forme $\langle B_i m_j, m_k B_l \rangle$ où les $B$ sont des états et les $m$ des mots.

### 8.5.3. *D'un problème à l'autre*

On peut remplacer la classe de questions

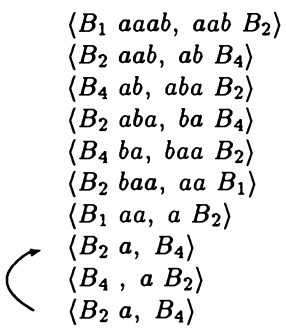



**Figure 8.17** *Machine de Post qui boucle*

telle machine de Post particulière s'arrêtera-t-elle pour $m = \sigma_1, \ldots, \sigma_n$ ?

par la classe de questions beaucoup plus générale :

> Pour une machine de Post $M$ quelconque définie sur l'alphabet $\Sigma = \{a, b\}$, et un mot $m$ formé de ces lettres par itération ($m \in \Sigma^*$), existe-t-il un algorithme qui détermine si $M$ s'arrête ou non pour $m$ ?

Ce problème a été démontré insoluble. De plus, on peut se servir de ce résultat pour établir l'insolubilité d'autres problèmes, par exemple, le problème de correspondance de Post, qui s'énonce comme suit :

> soient $A$ et $B$ deux listes de syllabes $\alpha$ et $\beta$ formées sur l'alphabet $\Sigma = \{0, 1\}$

(8.40) $$A = \alpha_1, \ldots, \alpha_k \quad \text{et} \quad B = \beta_1, \ldots, \beta_k;$$

> le problème est d'itérer, dans un certain ordre à déterminer, un certain nombre de fois, ces syllabes de manière à obtenir le même mot, en respectant la contrainte de *correspondance* : à chaque occurrence de $\alpha_i$ doit correspondre une occurrence de $\beta_i$.

*Exemple.* Donnons-nous l'ensemble de trois couples de syllabes

(8.41)
$$\begin{array}{ll} \alpha_1 = 1 & \beta_1 = 11 \\ \alpha_2 = 10 & \beta_2 = 0 \\ \alpha_3 = 11 & \beta_3 = 1 \end{array}$$

et demandons-nous si ce problème de correspondance a une solution. La réponse est affirmative :

(8.42) $$\langle \alpha_3, \alpha_1, \alpha_1, \alpha_2 \rangle = \langle \beta_3, \beta_1, \beta_1, \beta_2 \rangle,$$

(8.43) $$\langle 11, 1, 1, 10 \rangle = \langle 1, 11, 11, 0 \rangle.$$

Ce problème est une *instance* d'un problème général, le problème de correspondance pour deux suites équinombreuses de syllabes quelconques. Pour

prouver qu'il est insoluble, on utilise la technique de réduction exposée ci-dessous.

### 8.5.4. *La réduction*

On dit que la classe de problèmes $A$ se réduit à la classe de problèmes $B$ s'il existe une bijection qui associe à chaque instance du problème $A$ une instance du problème $B$, telle que la réponse à cette instance du problème $A$ soit "oui" si et seulement si la réponse à l'instance du problème $B$ est "oui". L'idée est celle-ci : on montre que le problème $B$ est *au moins aussi difficile* que le problème $A$ précédemment prouvé insoluble, et on en conclut que, *a fortiori*, le problème $B$ l'est aussi.

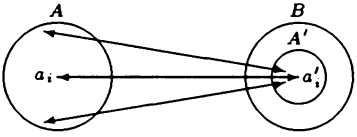



**Figure 8.18** *Réduction de A à B*

Il y a donc *bijection* entre l'ensemble des instances du problème $A$ et un *sous-ensemble propre* $A'$ de l'ensemble $B$. De ce que $A$ a été réduit à $B$, il ne découle nullement que les deux problèmes sont d'*égale* difficulté. Pour établir ce résultat, il faudrait non seulement réduire $A$ à $B$, mais aussi réduire $B$ à $A$.

### 8.5.5. *Réduction de l'arrêt à la correspondance*

En vue de prouver l'indécidabilité du problème de correspondance de Post, on va réduire à ce dernier le probléme de l'arrêt d'une machine de Post dont on sait qu'il est indécidable.

On réduit le problème de l'arrêt d'une machine de Post au problème de la correspondance de Post en montrant qu'il y a une bijection qui associe, à chaque instance $a_i$ du premier problème $A$, une instance $a'_i$ appartenant à un sous-ensemble $A'$ d'instances du deuxième problème $B$, telle que la réponse à la question $a_i$ est "oui" si et seulement si la réponse à la question $a'_i$ est "oui".

Pour cela, on commence par traduire les instructions de la machine de Post en couples de syllabes de la forme

(8.44) $\quad (\alpha_i, \beta_i)$.

A chaque mot que la machine de Post est capable de traiter (c'est-à-dire d'accepter ou de refuser), correspondra une solution au problème de Post formé à l'aide des lettres $\alpha$ et $\beta$. A chaque mot que la machine de Post est incapable

1. $\{B_1\}\ DO\ x = \phi \rightarrow \dot{A}CCEPT$ se réécrit $\langle *B * \phi, \phi \rangle$,
2. $(\{B_1\})\ \Box del(a, x) \rightarrow\ \{B_2\}\ IF\ x = \phi \rightarrow REJECT$ se réécrit
$$\langle *B_1 * a, B_2 * \rangle$$
$$\langle *B_2 * \phi, \phi \rangle$$
3. $(\{B_1\})\ \Box del(a, x) \rightarrow\ \{B_2\}; \Box del(a, x) \rightarrow\ \{B_4\}\ in(x, a)\{B_2\}$ se réécrit
$$\langle *B_1 * a, B_2 * \rangle$$
$$\langle *B_2 * a, B_4 * \rangle$$
$$\langle *B_4, a * B_2 * \rangle$$
4. $(\{B_1\})\ \Box del(a, x) \rightarrow\ \{B_2\}; \Box del(b, x) \rightarrow\ \{B_1\}$ se réécrit
$$\langle *B_1 * a, B_2 * \rangle$$
$$\langle *B_2 * b, B_1 * \rangle$$

Les trois autres constructions se traitent de manière analogue à 2-4.

**Figure 8.19** *Table de conversion*

de traiter (mot pour lequel elle boucle), correspondra un effort de solution au problème de correspondance indéfiniment reporté.

### 8.5.6. *Conversion des instructions*

On peut convertir en couples de syllabes les instructions d'une machine *quelconque* travaillant sur l'alphabet $\{a,b\}$. Ces couples de syllabes forment un problème de Post. la conversion se fait à l'aide des règles de réécriture suivantes.

### 8.5.7. *Exemple de conversion*

Si l'on applique ces directives générales à la machine de Post reproduite plus haut, on obtient le programme exprimé en couples à la figure 8.20. Par $\phi$, nous désignerons le mot vide dans la présente section. Nous ne distinguerons pas l'*acceptation* du *rejet*, qui sont chacune des instructions d'*arrêt*. Seul compte ici le contraste s'arrêter – boucler.

*Le mécanisme de progression de l'exécution*

Le système de Post est conçu de manière telle que le couple de départ est fixé d'avance : $i_1 = 0$ nécessairement, donc $\alpha_1 = B_0$ et $\beta_1 = B_0 * \sigma_1 * \ldots * \sigma_n * B_1$. La position de départ est décrite à la figure 8.21.

| Etiquettes | | $\langle \alpha\ ,\ \beta \rangle$ |
|---|---|---|
| $B_0$ | 1 | $\langle B_0\ ,\ B_0 * \sigma_1 \ldots \sigma_n * B_1 * \rangle$ |
| $B_1$ | 2 | $\langle * B_1 * a\ ,\ B_2 * \rangle$ |
| | 3 | $\langle * B_1 * b\ ,\ B_3 * \rangle$ |
| | 4 | $\langle * B_1 * \phi\ ,\ \phi \rangle$ |
| $B_2$ | 5 | $\langle * B_2 * a\ ,\ B_4 * \rangle$ |
| | 6 | $\langle * B_2 * b\ ,\ B_1 * \rangle$ |
| | 7 | $\langle * B_2 * \phi\ ,\ \phi \rangle$ |
| $B_3$ | 8 | $\langle * B_3 * a\ ,\ B_1 * \rangle$ |
| | 9 | $\langle * B_3 * b\ ,\ B_5 * \rangle$ |
| | 10 | $\langle * B_3 * \phi\ ,\ \phi \rangle$ |
| $B_4$ | 11 | $\langle * B_4\ ,\ a * B_2 * \rangle$ |
| $B_5$ | 12 | $\langle * B_5\ ,\ b * B_3 * \rangle$ |
| Boucle | 13 | $\langle * a\ ,\ a * \rangle$ |
| | 14 | $\langle * b\ ,\ b * \rangle$ |

**Figure 8.20** *Machine de Post avec instructions de recopiage*

$$\begin{array}{lcl} \alpha & : & B_0 \\ \beta & : & B_0 * \sigma_1 * \ldots * \sigma_n * B_1 \end{array}$$

**Figure 8.21** *Position de départ*

*Au départ*, $\beta$ est en avance puisque l'on a introduit le mot en $\beta$ ($\sigma_1 \ldots \sigma_n$ sont les places des lettres du mot). A l'étape suivante on recopie le mot en $\alpha$ et en $\beta$, ce qui maintient l'avance de $\beta$ sur $\alpha$ à cela près qu'un effacement de lettre peut avoir eu lieu ce qui diminue d'une unité l'avance de $\beta$ sur $\alpha$. la codification $\alpha$ rattrapera la codification $\beta$ si l'instance particulière du problème de correspondance admet une solution. A la suite de ce double recopiage, $B_1$

$$\begin{array}{lcl} \alpha & : & B_0 \boxed{* \sigma_1 * \ldots * \sigma_n * B_1} \\ & & \uparrow \quad \ldots \quad \uparrow \ \searrow \\ \beta & : & B_0\ * \sigma_1 * \ldots * \sigma_n * B_1\ \boxed{* \sigma_1 * \ldots * \sigma_n} \end{array}$$

**Figure 8.22** *Mécanisme de recopiage*

sera flanqué de tout le ruban à sa droite. On consulte alors le programme, où on trouve l'instruction

(8.45) $\langle \underbrace{B_1 * \sigma_1}_{\alpha}, \underbrace{B_i *}_{\beta} \rangle$.

On ajoute alors $B_1 * \sigma_1$ à la codification $\alpha$, et $B_i *$ à la codification $\beta$.

Après avoir fait les adjonctions précitées, on se trouve dans la situation suivante :

$$\begin{array}{lll} \text{Codification } \alpha & : & \overline{B_0 * \overbrace{\sigma_1 * \ldots * \sigma_n} * B_1 * \sigma_1}, \\ \text{Codification } \beta & : & B_0 * \sigma_1 * \ldots * \sigma_n * B_1 * \overbrace{\sigma_1 * \sigma_2 * \ldots \sigma_n} * B_i. \end{array}$$

**Figure 8.23** *Effet de recopiage*

C'est encore une situation de recopiage, mais comme en $\alpha$, la lettre $\sigma_1$ a été absorbée — ce qu'on représente en surmontant $B_1 * \sigma_1$ d'un trait continu $\overline{B_1 * \sigma_1}$ — le mot à recopier est désormais $w_1$, c'est-à-dire $w_0$ amputé de l'initiale $\sigma_1$.

L'état d'exécution est la suite $[B_0 w_0 \ldots B_j w_j \ldots B_n]$, où $B_j$ est la *j-ème étiquette* rencontrée dans l'exécution, et où $w_j$ est l'état du mot *avant* l'instruction si celle-ci est un test et *après* si c'est une affectation.

*Exemple.* Prenons comme entrée le mot *aab*. La machine de Post s'arrête (et rejette) ce mot au terme de la computation suivante : $\langle B_1 aab, abB_2\rangle$, $\langle B_2 ab, bB_4\rangle$, $\langle B_4 b, baB_2\rangle$, $\langle B_2 ba, aB_1\rangle$, $\langle B_1 a, B_2\rangle$, $\langle B_2\phi, \phi\rangle$, qui aboutit à un même mot (cf. figure 8.24).

*Remarque.* Se rappeler que dans le cas d'une affectation on ajoute, en $\beta$, $\sigma_i B_j$ et non simplement $B_j$.

Le calcul en vue de résoudre le problème de correspondance de Post construit à partir du programme de la machine et du mot-entrée suit exactement, dans sa codification $\beta$, les étapes du calcul du mot-entrée par la machine. On obtient deux occurrences d'un mot de 21 lettres.

$$\begin{array}{lll} \alpha & : & B_0 \overbrace{a\,a\,b}^{w_0} \overline{B_1\,a} \overbrace{a\,b}^{w_1} \overline{B_2\,a} \overbrace{b}^{w_2} \overline{B_4} \overbrace{b\,a}^{w_3} \overline{B_2\,b} \overbrace{a}^{w_4} \overline{B_1\,a}\, B_2 \overbrace{\phi}^{w_5} \\ \beta & : & \overline{B_0\,a\,a\,b\,B_1} \overbrace{a\,a\,b}^{w_0} \overline{B_2} \overbrace{a\,b}^{w_1} \overline{B_4} \overbrace{b}^{w_2} \overline{a\,B_2} \overbrace{b\,a}^{w_3} \overline{B_1} \overbrace{a}^{w_4} B_2 \overbrace{\phi}^{w_5} \end{array}$$

**Figure 8.24** *Exécution du programme avec l'entrée aab*

*Contre-épreuve*

Pour le mot *aa*, il y a une différence de deux unités entre les *b*, dont le nombre d'occurrences est 0, et les *a*. Dès lors, la machine boucle et les mots $\alpha$ et $\beta$ s'allongent indéfiniment, à tour de rôle (cf. figure 8.25).

$$\alpha \;:\; B_0 \overbrace{a\,a}^{w_0} \overline{B_1\,a} \overbrace{a}^{w_1} \overline{B_2\,a}\, B_4 \overbrace{a}^{w_2} \overline{B_2\,a}\, B_4 \overbrace{a}^{w_3} \ldots$$

$$\beta \;:\; \overline{B_0\,a\,a\,B_1} \overbrace{a\,a}^{w_0} \overline{B_2} \overbrace{a}^{w_1} B_4 \overbrace{a}^{w_2} \overline{B_2} \overbrace{a}^{w_3} B_4 \ldots$$

**Figure 8.25** *Système de Post dont la solution se dérobe indéfiniment*

## 8.6. Semi-décidabilité de la logique du premier ordre

Nous venons de montrer que le problème de l'arrêt d'une machine de Post est réductible au problème de correspondance de Post. Nous allons à présent montrer que le problème de correspondance de Post est réductible au problème de la validité dans le calcul des prédicats.

### 8.6.1. *Théorème de Floyd*

Pour établir la deuxième réduction qu'on vient de mentionner, il faut prouver qu'il existe une bijection qui associe, à *chaque* instance du problème de Post, *une* formule du calcul des prédicats telle que la réponse à la question "Cette instance du problème de correspondance de Post admet-elle une solution ?" est "oui" si et seulement si la réponse à la question "Cette formule du calcul des prédicats est-elle valide ?" est "oui". La présente section est reprise à Manna, 1974, p. 106.

Plus précisément encore, pour tout système de Post $S$ défini sur l'alphabet $\Sigma = \{a, b\}$, il existe une formule $W_S$ (de la forme 8.46) telle que $S$ est soluble si et seulement si $W_S$ est valide :

(8.46) $\quad (I \wedge II) \supset III,$

avec

$$\begin{array}{rcl} I & : & P(f_{\alpha_1}(c), f_{\beta_1}(c)) \wedge \ldots \wedge P(f_{\alpha_n}(c), f_{\beta_n}(c)), \\ II & : & \forall x \forall y\,(P(x,y) \supset (P(f_{\alpha_1}(x), f_{\beta_1}(y)) \wedge \ldots \wedge P(f_{\alpha_n}(x), f_{\beta_n}(y)))), \\ III & : & \exists z P(z,z). \end{array}$$

On démontre d'abord que si $W_S$ est valide, alors le système $S$ est soluble.

On suppose $\models W_S$, ce qui revient à supposer que $\|W_S\|_{\mathcal{D},\mathcal{I}} = \mathbf{V}$ pour tous les modèles $\langle \mathcal{D}, \mathcal{I} \rangle$. Adoptons pour domaine $\mathcal{D} = \{a, b\}^*$, c'est-à-dire $\{\phi, a, b, ab, aab, abb \ldots\}$. La fonction de désignation des constantes d'individus ($\mathcal{I}$) est définie par les équations

$$\begin{array}{lcll} \|c\|_{\mathcal{I}} & = & \phi & \text{mot vide,} \\ \|f_a x\|_{\mathcal{I}} & = & xa & \text{mot suivi de } a, \\ \|f_b x\|_{\mathcal{I}} & = & xb & \text{mot suivi de } b. \end{array}$$

La fonction d'évaluation des phrases (ouvertes) sera $\|P(x,y)\| = \mathbf{V}$ ssi $\exists \langle i_1, \ldots, i_m \rangle$ tel que $x = \alpha_{i_1} \ldots \alpha_{i_m} \wedge y = \beta_{i_1} \ldots \beta_{i_m}$.

Moyennant cette interprétation, l'antécédent $I$ de $W_S$ devient trivialement vrai. Remplaçons $P$, $f_{\alpha_i}$, $f_{\beta_i}$ et $c$ par leur valeur; nous obtenons $I'$ :

(8.47) $\phi\alpha_i = \alpha_i \wedge \phi\beta_i = \beta_i$,

pour chaque mot $x$ d'une syllabe, c'est-à-dire où $m = 1$.

*Remarque.* Le mot vide concaténé à un autre mot redonne ce mot : $\phi x = x = x\phi$.

Le deuxième antécédent ($II$) devient également trivialement vrai. L'interprétation de $II$ donne $II'$ :

(8.48) $x = \alpha_{i_1} \ldots \alpha_{i_m} \wedge y = \beta_{i_1} \ldots \beta_{i_m} \supset$
$x\alpha_i = \alpha_{i_1} \ldots \alpha_{i_{m+1}} \wedge y\beta_i = \beta_{i_1} \ldots \beta_{i_{m+1}}$.

C'est trivial si, sur base de l'hypothèse, on remplace $x$ et $y$ par leurs valeurs dans le conséquent de $II$. Ce conséquent devient

(8.49) $\alpha_{i_1} \ldots \alpha_{i_m}\alpha_i = \alpha_{i_1} \ldots \alpha_{i_{m+1}} \wedge \beta_{i_1} \ldots \beta_{i_m}\beta_i = \beta_{i_1} \ldots \beta_{i_{m+1}}$.

Comme $(I \wedge II) \supset III$ est valide, par hypothèse, et que $\|I\| = \mathbf{V}$ et $\|II\| = \mathbf{V}$, on peut en conclure que, sous cette même interprétation, on a $\|III\| = \mathbf{V}$, donc que $\exists z(z = \alpha_{i_1} \ldots \alpha_{i_m} \wedge z = \beta_{i_1} \ldots \beta_{i_m})$. Deux quantités égales à une même troisième ($z$) étant égales entre elles, on a $\alpha_{i_1} \ldots \alpha_{i_m} = \beta_{i_1} \ldots \beta_{i_m}$, et dès lors le problème de correspondance de Post a une solution, donc

(8.50) Si $\models W_S$ alors $R(S)$.

On démontre ensuite la réciproque.

On suppose $R(S)$, et on cherche à prouver que $\models (I \wedge II) \supset III$. Pour cela, il suffit de montrer que toute interprétation qui rend vrais $I$ et $II$ rend vrai $III$, compte tenu de l'hypothèse que $R(S)$.

Voici le schéma de la démonstration : supposons $I'$ vrai, c'est-à-dire

(8.51) $P(f_{\alpha_i}(c), f_{\beta_i}(c))$,

supposons $II'$ vrai, c'est-à-dire

(8.52) $P(x, y) \supset P(f_{\alpha_i}(x), f_{\beta_i}(y))$;

remplaçons $x$ par $f_{\alpha_{i_1}}(c)$ et $y$ par $f_{\beta_{i_1}}(c)$ :

(8.53) $\underbrace{P(f_{\alpha_{i_1}}(c), f_{\beta_{i_1}}(c))}_{I'} \supset \underbrace{P(f_{\alpha_{i_2}}(f_{\alpha_{i_1}}(c)), f_{\beta_{i_2}}(f_{\beta_{i_1}}(c)))}_{II'}$.

On peut donc détacher $II'$ et recommencer. A cette fin, on remplace $x$ dans $I'$ par $f_{\alpha_{i_2}}(f_{\alpha_{i_1}}(c))$ et $y$ par $f_{\beta_{i_2}}(f_{\beta_{i_1}}(c))$, et ainsi de suite. Les propositions $I$ et $II$ combinées font ainsi office de schéma d'induction.

On applique en fin de parcours la convention abbréviative remplaçant

(8.54) $f_{\sigma_m}(\cdots(f_{\sigma_2}(f_{\sigma_1}(x)))\cdots)$,

où $\sigma_i \in \{a, b\}$, par $f_{\sigma_1\sigma_2\ldots\sigma_m}(x)$. $II'$ devient alors $II''$ :

(8.55) $P(f_{\alpha_{i_1}\ldots\alpha_{i_m}}, f_{\beta i_1\ldots\beta_{i_m}})$.

Or, en vertu de l'hypothèse $R(S)$, on sait que $\alpha_{i_1} \ldots \alpha_{i_m} = \beta i_1 \ldots \beta_{i_m}$. On peut donc remplacer les $\alpha_{i_j}$ et les $\beta_{i_j}$ par des $\gamma_{i_j}$, et on obtient successivement

$$
\begin{array}{ll}
(8.56) & P(\underbrace{f_{\gamma_{i_1}\ldots\gamma_{i_m}}}, \underbrace{f_{\gamma_{i_1}\ldots\gamma_{i_m}}}), \\
& P(\quad z \quad , \quad z \quad), \\
\exists z & P(\quad z \quad , \quad z \quad).
\end{array}
$$

### 8.6.2. *Théorème de Church*

**Théorème 8.7 : théorème de Church.** Le calcul des prédicats est *indécidable*, ou plus exactement *semi-décidable*. En effet, les schémas valides peuvent être découverts mécaniquement. Leur ensemble est récursivement énumérable. Mais les schémas non-valides ({schémas simplement consistants} ∪ {schémas inconsistants}) constituent un ensemble qui n'est pas récursivement énumérable. En revanche, l'ensemble des schémas inconsistants est récursivement énumérable.

*Démonstration.* Si le calcul des prédicats était décidable, le problème de correspondance de Post le serait aussi, et si celui-ci l'était, le problème de l'arrêt le serait également. Or le problème de l'arrêt est insoluble par algorithme, donc le calcul des prédicats est indécidable (cf. figure 8.26).

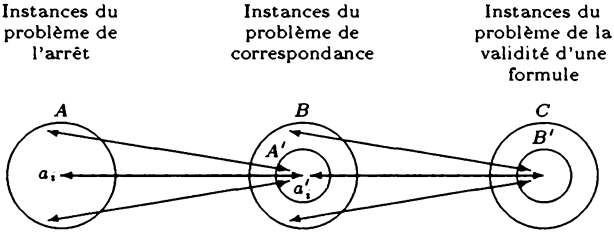



**Figure 8.26** *Réduction de A à C*

## 8.7. Du calcul des prédicats aux théories du premier ordre

On obtient une théorie du premier ordre en sélectionnant certaines constantes du calcul des prédicats et en ajoutant aux axiomes logiques des postulats spécifiques, comme par exemple, les axiomes de Peano.

Les recherches en théorie de la preuve comme celles qui ont révélé la possibilité de *normaliser* les preuves en déduction naturelle ou d'*éliminer la coupure* en calcul des séquents gardent tout leur intérêt quand on passe du calcul des prédicats aux théories qui prennent appui sur celui-ci. Elles acquièrent même une certaine utilité pratique :

> Du point de vue de l'informatique, un champ [...] intéressant d'application pour les algorithmes de normalisation est la possibilité d'extraire le contenu constructif d'un raisonnement mathématique complexe. De tels algorithmes peuvent fournir des programmes vérifiés à partir d'une dérivation et prouver que certaines spécifications sont remplies (Schwichtenberg, à paraître).

Nous avons signalé au début du chapitre 6 que l'introduction de l'induction mathématique dans un système axiomatique exerçait une action discriminatoire sur les méthodes de logique en empêchant ou en limitant la possibilité de normalisation des preuves et d'élimination de la coupure, mettant fin du même coup à l'interchangeabilité entre méthodes synthétiques et méthodes analytiques de preuve.

Cette incompatibilité entre induction et normalisation a conduit à des compromis intéressants : on se limite à un sous-système de l'arithmétique de Peano dans lequel on restreint la règle d'induction mathématique aux formules existentielles de la forme $A(x)$ où

(8.57) $\quad A(x) \equiv \exists z_1, \ldots, \exists z_n B(x, z_1, \ldots, z_n)$

Dans ce sous-système, on peut éliminer la coupure jusqu'à ce qu'on arrive à des formules du type de la formule précitée.

Que gagne-t-on à restreindre de la sorte ses moyens de preuve ? La réponse tient dans le résultat suivant dû à Kreisel, Parsons, Mints et d'autres :

> Si une formule de la forme
>
> (8.58) $\quad \forall x \exists z B(x, z)$
>
> est démontrable dans ce sous-système de l'arithmétique, alors il existe une fonction récursive primitive $f$ telle que la formule
>
> (8.59) $\quad \forall x B(x, f(x))$
>
> est vraie dans $\mathcal{N}$.

## 8.8. La ligne de démarcation entre logique et mathématique

La thèse reçue aujourd'hui, est que le contenu proprement mathématique d'une théorie, comme l'arithmétique ou la géométrie par exemple, se concentre dans les axiomes propres de la théorie et dans les théorèmes qui en découlent, la relation d'implication existant entre les axiomes et les théorèmes relevant de la pure logique.

Cette thèse est formulée avec clarté par Lewis et Langford :

> Si la vérité essentielle affirmée par les mathématiques est que certaines suppositions [les postulats] impliquent certains théorèmes, et si chaque implication en mathématique est une instance de quelque principe universellement valide de déduction (principe de logique), alors il ne peut

> y avoir dans une preuve mathématique aucune étape qui dépende, par exemple, du caractère particulier de notre espace ou des propriétés empiriques des collections que l'on peut compter (Lewis et Langford, 1932, 1959, p. 21).

Développant un exemple de Beth, nous avons montré en 5.3.9 comment une démonstration de géométrie qui selon Kant nécessitait le recours à l'intuition de l'espace pouvait être réécrite de manière telle que la dérivation soit purement logique. On aurait pu choisir un exemple arithmétique : dériver par les seules règles du calcul des prédicats la commutativité de la somme arithmétique ($x + y = y + x$) à partir des axiomes de Peano ou d'une variante de ce système axiomatique (Mendelson, 1965, p. 105).

Le fait que la présence de l'induction mathématique dans une théorie empêche ou limite, au moins dans certains cas, la normalisation des preuves ou l'élimination de la coupure semble cependant indiquer que la conception classique de la séparation entre la *forme logique* et le *contenu mathématique* d'une théorie formalisée n'est pas aussi nette qu'on l'a cru. Le contenu semble altérer la forme.

# Chapitre 9
# Théories du premier ordre

## 9.1. Introduction

Nous avons vu au chapitre précédent que le calcul des prédicats permet de représenter des concepts, ainsi que les raisonnements effectués à leur sujet. Le calcul des prédicats est universel, en ce sens que ses axiomes et ses règles d'inférence sont communs à toutes les théories scientifiques et que son lexique extralogique (constantes prédicatives et fonctionnelles) peut exprimer les constantes de toutes les théories. Il ne permet cependant pas une représentation naturelle de tous les raisonnements valides; pour cette raison, on enrichit parfois son lexique logique, notamment d'opérateurs modaux (voir volume II).

Bien souvent, cette universalité n'est pas souhaitée. L'objectif habituel d'un utilisateur de la logique est l'étude rigoureuse d'un domaine précis. Les exemples les plus connus appartiennent au domaine des mathématiques. Cela ne signifie pas que seuls les mathématiciens peuvent utiliser la logique avec profit; simplement, quand un certain concept a reçu un modèle formel, il devient par le fait même un objet mathématique, qu'il ait ou non une utilité concrète. Un exemple typique est celui de programme d'ordinateur.

Dans ce chapitre, nous présentons le moyen le plus direct d'utiliser la logique dans un domaine particulier. Le calcul des prédicats est à la fois un langage et un outil de raisonnement; ces deux aspects peuvent être adaptés à un domaine précis. Le langage du calcul des prédicats comporte, en plus des symboles purement logiques (connecteurs, quantificateurs et variables), des constantes prédicatives et fonctionnelles, en nombre illimité et susceptibles de recevoir les interprétations les plus variées. L'adaptation du langage consistera en la sélection des constantes nécessaires à l'expression des concepts pertinents au

domaine visé. L'outil de raisonnement ne doit pas être modifié mais, en général, certaines relations existent entre les concepts modélisés par les constantes du langage. Ces relations seront elles-mêmes formalisées par des axiomes spéciaux, appelés *postulats*.

Considérons deux exemples bien connus de mathématique élémentaire. Dans l'ensemble $\mathbf{Z}$ des entiers relatifs[1], l'addition est une opération associative, c'est-à-dire que $(a+b)+c = a+(b+c)$; l'entier 0 est neutre pour cette opération $(a+0 = a = 0+a)$ et chaque élément $a$ possède un opposé $-a$, tel que $a+(-a) = 0 = (-a)+a$. Une situation analogue existe pour la multiplication dans l'ensemble $\mathbf{R_0}$ des réels[2] non nuls : la multiplication est associative et admet le neutre 1; de plus, chaque élément $a$ admet l'inverse $1/a$, tel que $a.(1/a) = 1 = (1/a).a$. Pour souligner l'analogie, on qualifie de *groupes* les structures telles que $(\mathbf{Z}, +)$ et $(\mathbf{R_0}, .)$. Les groupes se rencontrent dans toutes les branches des mathématiques et, pour cette raison, il est intéressant et utile de définir une théorie des groupes. Un groupe[3] est un ensemble, muni d'une opération binaire interne et partout définie. L'opération est associative et admet un élément neutre; chaque élément du groupe admet un élément symétrique.

Un langage adéquat pour cette théorie comportera nécessairement une constante fonctionnelle à deux arguments pour modéliser l'opération; de même, une constante individuelle représentera l'élément neutre. Classiquement, la constante fonctionnelle est notée par un point (utilisé en notation infixée) et l'élément neutre est représenté par $e$. Les propriétés de ces deux objets sont modélisées par les postulats A1 à A3 :

(9.1)
$$\begin{aligned} &\text{A1} : \forall x \forall y \forall z\,((x.y).z = x.(y.z))\,, \\ &\text{A2} : \forall x\,(x.e = x = e.x)\,, \\ &\text{A3} : \forall x \exists y\,(x.y = e = y.x)\,. \end{aligned}$$

Le premier traduit l'associativité, le deuxième le fait que $e$ désigne l'élément neutre et le troisième l'existence, pour tout élément, d'un élément symétrique. (Le caractère interne et partout défini de l'opération ne fait pas l'objet d'un axiome spécial; en effet, une constante fonctionnelle à $n$ arguments ne peut être interprétée que comme une opération $n$-aire interne et partout définie.)

L'entité formée du langage et des postulats est, au sens formel, une théorie. Il convient, à propos de cet exemple, d'observer une série de faits.

1. La conformité du modèle à la réalité que l'utilisateur souhaite représenter n'est pas garantie; elle n'est pas du ressort de la logique. D'ailleurs, la

[1] $\mathbf{Z} = \{0, 1, -1, 2, -2, \ldots\}$

[2] Les réels ont étés introduits pour "boucher des trous" existant dans l'ensemble $\mathbf{Q}$ des nombres rationnels.

[3] Nous donnons ici la définition traditionnelle. En théorie des catégories, un groupe se définit comme une catégorie dont tous les morphismes sont inversibles et qui n'a qu'un seul objet.

réalité peut n'être qu'imparfaitement cernée par l'utilisateur, alors que le modèle a une signification précise.

2. Le passage d'une réalité à son modèle implique certains choix, parfois inconscients. Par exemple, nous avons imposé à la constante $e$ d'être neutre à gauche et à droite, mais nous n'avons pas explicitement imposé son unicité. De même, rien n'est dit de l'unicité du symétrique d'un élément donné. Il va de soi que ces choix ne sont généralement pas sans conséquence. Nous démontrerons cependant que l'unicité du neutre et du symétrique est une conséquence logique des postulats A1 à A3.

3. Une même réalité peut se traduire de plusieurs manières. Puisque chaque élément admet un et un seul symétrique (ce fait sera prouvé plus loin), on pourrait adjoindre une constante fonctionnelle $i$ au langage, $i(x)$ se lisant "le symétrique de $x$".

4. Une même réalité peut être modélisée de manière plus ou moins fine. La théorie que nous venons d'introduire rend bien compte de ce qui est un groupe et de ce qui ne l'est pas; elle permet aussi d'exprimer des propriétés communes à tous les groupes. Le langage permet l'expression de faits vrais dans certains groupes et faux dans d'autres (la commutativité, par exemple). Toutefois, certaines notions relatives au concept de groupe ne sont pas représentables dans le langage (le fait qu'il s'agisse d'un groupe fini, par exemple).

5. Bien souvent, on modélise un domaine sans le connaître parfaitement; on souhaite d'ailleurs que la modélisation apporte des renseignements nouveaux à propos du domaine. Cette attitude peut être poussée à l'extrême, et une théorie formelle peut être donnée a priori, sans référence à une quelconque réalité.

6. Signalons enfin que le calcul des prédicats, ainsi que la théorie des groupes et les autres théories dérivées du calcul des prédicats sont généralement dits "du premier ordre". Cette expression rappelle que, dans toute interprétation, les variables correspondent à des éléments du domaine d'interprétation, et non à des objets complexes, tels les sous-ensembles de ce domaine, ou les fonctions définies sur lui. En logique comme en informatique, on peut attribuer à un objet un *type*, qui détermine les opérations que l'on peut effectuer sur cet objet. Les objets sont habituellement définis de manière structurée, c'est-à-dire qu'un objet complexe est constitué de sous-objets plus simples. Ceci donne lieu à une hiérarchie de types, dont l'étude est le but de la théorie des types. La logique du premier ordre fait abstraction des types, ce qui lui confère une certaine simplicité. Dans ce contexte, les objets structurés s'étudient principalement dans le cadre de la théorie des ensembles (du premier ordre). Si on souhaite modéliser explicitement le type des objets, on doit considérer des théories d'ordre supérieur. Les deux points de vue ont leur intérêt,

et seront abordés dans le second volume.

Le concept de théorie formelle est habituellement développé dans deux directions. D'une part, le problème de la correspondance existant entre une théorie et ses modèles est étudié dans un cadre très général. Cette étude débouche sur des méthodes systématiques de construction de modèles d'une théorie donnée; la plupart de ces méthodes font appel à des notions non élémentaires d'algèbre. D'autre part, certaines théories particulières font l'objet d'une étude approfondie, parce que leurs modèles ont une importance cruciale en mathématique. Le plus souvent, cette seconde démarche nécessite, elle aussi, le recours à des concepts mathématiques difficiles.

L'approche adoptée dans ce chapitre est différente. Les concepts généraux les plus importants sont introduits par des exemples, ce qui facilite leur compréhension. Deux familles de théories seront abordées dans la suite du chapitre; la première concerne la notion de relation d'ordre, tandis que la seconde se rapporte aux structures algébriques les plus courantes (groupe, anneau, corps et champ). Ces notions sont bien sûr importantes en mathématique; en outre, les théories correspondantes permettent d'introduire et d'illustrer les notions importantes de décidabilité (2.4.5), de complétude (7.7.2 et 7.7.3) et de catégoricité (7.6.2).

Deux théories ont, en mathématique et en informatique, une importance particulière : la théorie des ensembles et la théorie des nombres. La première sera abordée dans le chapitre suivant (volume II).

## 9.2. Quelques définitions

### 9.2.1. *Généralités*

Il est commode de répartir les symboles du langage des prédicats en deux catégories. Les symboles *logiques* sont les connecteurs (y compris l'égalité), les quantificateurs, les variables et les séparateurs (parenthèses, crochets, accolades, etc.). Les symboles *extralogiques* sont les constantes fonctionnelles et les constantes prédicatives. Dans le calcul des prédicats, l'ensemble des symboles extralogiques est illimité; l'interprétation de ces symboles n'est pas fixée. Une *signature* est un sous-ensemble de l'ensemble des symboles extralogiques. A toute signature est associé un *langage*, qui est l'ensemble des formules du calcul des prédicats dont tous les symboles extralogiques appartiennent à la signature.

Le langage de la théorie des groupes évoquée plus haut est associé à la signature composée de la constante individuelle $e$ et de la constante fonctionnelle à deux places symbolisée par un point. La plupart des théories "intéressantes" ont pour signature un ensemble fini relativement petit.

Une *théorie* se compose d'un langage, défini par une signature, et d'un ensemble récursif de formules du langage, appelées *axiomes* ou *postulats*. Les

conséquences logiques des postulats sont appelées *théorèmes*. Supposons que l'on ait déterminé un système axiomatique adéquat et complet du calcul des prédicats; une formule du langage sera un théorème si et seulement si le système permet de la déduire des axiomes. Le nom de théorie est tantôt attribué à l'ensemble des postulats, tantôt à l'ensemble des théorèmes. Dans la deuxième acception, on appelle *théorie* le plus petit ensemble de formules contenant les postulats et fermé sur la relation de dérivabilité. Cette diversité d'acception du mot "théorie" n'induit généralement pas d'ambiguïté.

On peut toujours remplacer un postulat contenant des variables libres par la fermeture universelle de celui-ci, sans modifier l'ensemble des conséquences logiques de la théorie. Dans la suite, on supposera que les postulats sont des formules fermées.

L'acception formelle du mot "théorie" est voisine de son acception usuelle. L'idée, en effet, est de considérer les postulats comme des vérités premières, et d'en étudier les conséquences. L'intérêt de cette démarche est que, bien souvent, les postulats paraissent "évidents", alors que certaines de leurs conséquences ne le sont pas. La géométrie classique est un bon exemple de cette situation. Dans les sciences expérimentales, on peut rencontrer la situation inverse. Une série de faits expérimentaux sont connus, et l'on cherche à les justifier, c'est-à-dire à les déduire d'un ensemble de prémisses, baptisées lois ou postulats. La qualité d'un ensemble de postulats est liée à la quantité de faits qu'il permet d'expliquer. Ces faits sont souvent mesurables : de la qualité de l'ensemble des postulats dépend aussi le degré de concordance entre les mesures prévues et les mesures observées. Les développements de l'astronomie de position, de Ptolémée à Einstein, en passant par Copernic, Kepler et Newton, illustrent bien la notion de théorie, telle qu'elle est utilisée dans les sciences naturelles. La théorie d'Einstein est meilleure que celle de Newton parce que, notamment, elle prédit avec plus d'exactitude les temps et lieux des passages de Mercure au périhélie. Dans le cas des théories formelles, la situation est plus simple, car la mesure de l'écart entre une proposition et un ensemble de postulats ne peut prendre que deux valeurs (cf. chapitre 7 : une proposition *est* ou *n'est pas* conséquence logique d'un ensemble de postulats, sans possibilité intermédiaire).

Dans le cadre du premier ordre, on sait que si une formule $A$ est une conséquence logique d'un ensemble $E$ de postulats, ce fait pourra toujours être démontré, en utilisant une méthode de déduction adéquate et complète (la méthode de résolution, ou celle de déduction naturelle, par exemple). Un problème fort différent, souvent plus délicat, consiste à prouver qu'une formule $B$ n'est *pas* conséquence logique d'un ensemble $E$ de postulats. La méthode habituelle consiste à construire un modèle de l'ensemble $E$ dans lequel la formule $B$ est fausse. Par exemple, pour démontrer que la formule $\forall x \forall y (x.y = y.x)$ n'est pas un théorème de la théorie des groupes, il suffit d'exhiber un groupe non abélien, c'est-à-dire un groupe dont l'opération n'est pas commutative.

Cette technique de preuve par contre-exemple est d'emploi courant, même en dehors de la logique formelle. Un exemple célèbre est celui de la géométrie plane classique. Si l'on veut montrer que le "postulat d'Euclide" n'est pas conséquence logique des autres postulats de la géométrie, il suffit de construire une géométrie non-euclidienne, c'est-à-dire une "structure" (une entité jouant le rôle du plan euclidien) dans laquelle tous les postulats de la géométrie sont vrais, à l'exception de ce postulat. La découverte de telles géométries, notamment par Bolyai et Lobatchevsky, vers 1830, a fort surpris la communauté des mathématiciens. Cet exemple maintenant classique est intéressant, parce qu'il met en évidence un phénomène fréquent. Quand on axiomatise une certaine réalité, on ne rend généralement compte que d'une partie des propriétés de cette réalité. Cela implique souvent que la théorie admet des modèles "non réguliers", c'est-à-dire des modèles qui diffèrent de la réalité appréhendée. On peut essayer d'éliminer les modèles non réguliers (*non intended*) par l'adjonction de nouveaux postulats mais, comme nous le verrons plus loin, cette approche a des limites.

Une *structure* relative à un langage $\mathcal{L}$ est une interprétation de ce langage. Une structure comporte donc un domaine $\mathcal{D}$; de plus, pour chaque constante fonctionnelle ou prédicative appartenant à la signature du langage, la structure comporte une application de $\mathcal{D}^n$ dans $\mathcal{D}$ (pour les symboles de fonction) ou une application de $\mathcal{D}^n$ dans $\{\mathbf{V},\mathbf{F}\}$ (pour les prédicats). Rappelons que l'on appelle *structure* $(\mathcal{S})$ le couple $\langle \mathcal{D}, \mathcal{I} \rangle$ qui sert à *interpréter* les expressions du langage et qu'on réserve le terme modèle $\mathcal{M}$ aux structures qui *rendent vraies* les axiomes et les théorèmes d'une théorie.

Soit $\mathcal{S}$ une structure de domaine $\mathcal{D}$ et soit $\mathcal{A}$ une assignation aux variables dans $\mathcal{D}$ (ce que nous avons précédemment appelé une assignation $\mathcal{A}$). Si $\phi(x_1, \ldots, x_n)$ est une formule de $\mathcal{L}$ ne comportant pas d'autres variables libres que $x_1, \ldots, x_n$, alors $\mathcal{S}$ et $\mathcal{A}$ associent à la formule une valeur de vérité. On écrira $\mathcal{S}, \mathcal{A} \vDash \phi(x_1, \ldots, x_n)$ si et seulement si cette valeur est V. Si $d_1, \ldots, d_n$ sont respectivement les images de $x_1, \ldots, x_n$ par $\mathcal{A}$, on dira plus simplement que $\phi(d_1, \ldots, d_n)$ est vrai dans $\mathcal{S}$. Rappelons que la formule $\phi(d_1, \ldots, d_n)$ (qui n'est plus une formule de $\mathcal{L}$ car $\mathcal{L}$ ne contient pas de constantes individuelles $d_i$) est obtenue en remplaçant dans $\phi(x_1, \ldots, x_n)$ toute occurrence libre de $x_i$ par $d_i$, pour $i = 1, \ldots, n$.

Une structure et une interprétation des variables déterminent donc une partition des formules de $\mathcal{L}$ en *formules vraies* et *formules fausses*. En particulier, une structure détermine une partition des formules fermées du langage en formules vraies et formules fausses. Une structure est un *modèle* d'une théorie si elle rend vrais tous les postulats (et du même coup tous les théorèmes) de la théorie.

Toute théorie de langage $\mathcal{L}$ détermine une partition de l'ensemble des structures de langage $\mathcal{L}$ en deux sous-ensembles : les modèles, et les autres structures. Par exemple, $(\mathbf{Z}, +, 0)$ et $(\mathbf{N}, +, 0)$ sont des structures relatives au lan-

gage de la théorie des groupes; la première est un modèle de cette théorie mais la seconde ne l'est pas (le postulat d'existence du symétrique n'est pas vérifié).

Une théorie $T$ de langage $\mathcal{L}$ partitionne les formules de $\mathcal{L}$ en trois sous-ensembles. Une formule $A$ est *$T$-valide* si tous les modèles de $T$ la rendent vraie; elle est *$T$-inconsistante* si $\neg A$ est $T$-valide, c'est-à-dire si tous les modèles de $T$ rendent fausse la formule $A$. Enfin, cette formule est *$T$-contingente* si elle n'est ni $T$-valide, ni $T$-inconsistante, c'est-à-dire si elle est rendue vraie par certains modèles de $T$ et fausse par d'autres. (Une formule est *$T$-consistante* si elle n'est pas $T$-inconsistante.) Ces classifications respectivement trichotomique et dichotomique sont résumées dans la figure 9.1.

| $T$-valide | $T$-contingente[4] | $T$-inconsistante |
|---|---|---|
| $T$-consistante | | $T$-inconsistante |

**Figure 9.1** *Classification des formules par rapport à une théorie*

Le calcul des prédicats peut être vu comme une théorie du premier ordre particulière. Sa signature comporte un nombre illimité de constantes fonctionnelles et prédicatives, tandis que l'ensemble des postulats est vide. La théorie dont la signature et l'ensemble des postulats sont tous les deux vides est la théorie de l'égalité pure. En effet, cette théorie ne contient aucune constante extralogique et aucun postulat, la constante "=" appartenant au calcul des prédicats avec identité[5].

Une théorie se compose d'un langage et de postulats appartenant à ce langage. Il est donc indispensable de pouvoir reconnaître les formules du langage, ce qui suppose que la signature soit un ensemble récursif de symboles. Le langage et l'ensemble des postulats sont égalements récursifs.

### 9.2.2. *Résultats élémentaires de la théorie des groupes*

A titre de premier exemple, nous démontrons ici quelques théorèmes élémentaires de la théorie des groupes. Rappelons d'abord qu'il existe plusieurs manières d'axiomatiser la notion de groupe. Le choix de la signature ne détermine nullement le choix des postulats. Tout d'abord, il est toujours possible de modifier un ensemble de postulats de manière non essentielle, par l'adjonction de conséquences logiques de postulats précédemment introduits, voire même de simples tautologies. Ce type de modification est sans utilité; au

[4] Nous appelons $T$-contingente ce que nous avons appelé précédemment formule simplement consistante, c'est-à-dire consistante mais non valide. La terminologie n'est pas fixée.

[5] Les propriétés de l'égalité, souvent désignées sous le nom de règles de Leibniz (1.5.2), font partie intégrante du calcul des prédicats avec identité.

contraire, on souhaite parfois vérifier que tout postulat est indispensable, c'est-à-dire qu'aucun postulat n'est conséquence logique des autres. Pour prouver qu'un énoncé $A$ est *indépendant* d'un ensemble $E$ d'énoncés, il suffit d'exhiber un modèle de $E$ dans lequel $A$ est faux. Au paragraphe précédent, on a mentionné le fait que le postulat d'Euclide était indépendant des autres postulats de la géométrie. Nous démontrons à présent que chaque postulat de la théorie des groupes est indépendant des autres.

Considérons d'abord la structure (N, +, 0). L'addition des entiers naturels est bien associative et admet le neutre 0; par contre, les entiers naturels autres que 0 n'admettent pas (dans N) d'élément symétrique pour l'addition. D'autre part, la structure (N, &, 0), où & désigne l'opération triviale appliquant tout couple d'entiers naturels sur 0, vérifie le postulat d'associativité A1 et le postulat d'existence du symétrique A3, avec $e = 0$ (on a en effet $\forall x\,(x\&0 = 0 = 0\&x)$, dont la formule $\forall x\,\exists y\,(x\&y = 0 = y\&x)$ est conséquence logique), mais ne vérifie pas le postulat d'existence du neutre A2 (c'est-à-dire $\forall x\,(x.e = x = e.x)$). Enfin, considérons la structure $(\{0, 1, \infty\}, *, 1)$, où $*$ désigne la "multiplication étendue", définie par la table présentée dans la figure 9.2[6]

| $*$ | 0 | 1 | $\infty$ |
|---|---|---|---|
| 0 | 0 | 0 | 1 |
| 1 | 0 | 1 | $\infty$ |
| $\infty$ | 1 | $\infty$ | $\infty$ |

**Figure 9.2** *Table de multiplication étendue*

On vérifie aisément que 1 est neutre, que chaque élément admet un symétrique, mais aussi que l'opération $*$ n'est pas associative : on a

(9.2) $\quad (\infty * \infty) * 0 = 1$

et

(9.3) $\quad \infty * (\infty * 0) = \infty\,.$

Le deuxième axiome de la théorie des groupes donne une propriété de l'élément neutre $e$. Il est intéressant de montrer que cette propriété caractérise $e$, c'est-à-dire que seul l'élément $e$ vérifie la propriété. Formellement, cela revient à prouver la formule

(9.4) $\quad \forall a\,(\forall x\,(x.a = x = a.x) \supset a = e)\,.$

Voici un schéma de preuve :

[6] Le symbole $\infty$ n'a ici aucune signification particulière mais la définition de l'opération $*$ permet, dans une certaine mesure, d'interpréter ce symbole comme "l'infini".

| | | |
|---|---|---|
| 1. | $\forall x\,(x.a = x = a.x)$ | Hypothèse |
| 2. | $\quad e.a = e$ | 1 |
| 3. | $\quad a = e$ | 2, A2 |
| 4. | $\forall x\,(x.a = x = a.x) \supset (a = e)$ | 3, Décharge 1 |
| 5. | $\forall a\,(\forall x\,(x.a = x = a.x) \supset (a = e))$ | 4 |

La formule

(9.5) $\quad \forall x \forall y \forall z\,((x.z = y.z) \supset (x = y))$

est un théorème de la théorie des groupes; elle exprime la possibilité de simplifier les égalités. Voici un schéma de preuve (les axiomes A1, A2, A3 sont donnés en 9.1) que le lecteur pourra compléter :

| | | |
|---|---|---|
| 1. | $x.z = y.z$ | Hypothèse |
| 2. | $\quad z.a = e$ | Hypothèse |
| 3. | $\quad\quad (x.z).a = (y.z).a$ | 1 |
| 4. | $\quad\quad x.(z.a) = y.(z.a)$ | 3, A1 |
| 5. | $\quad\quad x.e = y.e$ | 2, 4 |
| 6. | $\quad\quad x = y$ | 5, A2 |
| 7. | $\quad (z.a = e) \supset x = y$ | 6, Décharge 2 |
| 8. | $\quad \exists u\,(z.u = e) \supset x = y$ | 7 |
| 9. | $\quad x = y$ | 8, A3 |
| 10. | $(x.z = y.z) \supset (x = y)$ | 9, Décharge 1 |
| 11. | $\forall x \forall y \forall z\,((x.z = y.z) \supset (x = y))$ | 10 |

La formule analogue $\forall x \forall y \forall z\,((z.x = z.y) \supset (x = y))$ se prouve de la même manière. A partir de ce résultat, on établit aussi que $x.y = e$ si et seulement si $y.x = e$; en effet, dans le tableau suivant, les formules 1, 2, 4 et 5 sont équivalentes.

1. $x.y = e$
2. $(x.y).x = e.x$
3. $e.x = x.e$
4. $x.(y.x) = x.e$
5. $y.x = e$

Le troisième axiome de la théorie des groupes affirme l'existence, pour tout élément $x$, d'un symétrique $y$ tel que $x.y = e = y.x$. Ici aussi, il est intéressant de prouver l'unicité du concept de symétrique (cela justifiera la notation $x^{-1}$, souvent utilisée pour désigner le symétrique de $x$). L'unicité du symétrique est un corollaire immédiat des résultats précédents. Pour tout $x$, il existe un et un seul élément, noté habituellement $x^{-1}$, tel que $x.x^{-1} = e = x^{-1}.x$; de plus, pour tout $y \neq x^{-1}$, on a $x.y \neq e \wedge y.x \neq e$.

### 9.2.3. *Applications*

Les notions très simples introduites dans le début de ce chapitre permettent déjà de présenter formellement divers problèmes et leurs solutions. Nous restons dans le cadre de la théorie des groupes, dont la richesse ne doit pas être masquée par la simplicité et le petit nombre des axiomes.

#### 9.2.3.1. *Unicité du symétrique : preuve directe*

Une particularité notable des systèmes axiomatiques usuels est leur non-déterminisme : un même théorème peut être établi de plusieurs façons différentes. A titre d'exemple, voilà une preuve directe de l'unicité du symétrique :

| | | |
|---|---|---|
| 1. | $y.x = e$ | Hypothèse |
| 2. | $\quad x.z = e$ | Hypothèse |
| 3. | $\quad\quad (y.x).z = e.z$ | 1 |
| 4. | $\quad\quad e.z = z$ | A2 |
| 5. | $\quad\quad (y.x).z = z$ | 3, 4 |
| 6. | $\quad\quad y.(x.z) = y.e$ | 2 |
| 7. | $\quad\quad y.e = y$ | A2 |
| 8. | $\quad\quad y.(x.z) = y$ | 6,7 |
| 9. | $\quad\quad (y.x).z = y.(x.z)$ | A1 |
| 10. | $\quad\quad y = z$ | 5,8,9 |
| 11. | $\quad (x.z = e) \supset (y = z)$ | 10, Décharge 2 |
| 12. | $(y.x = e \wedge x.z = e) \supset (y = z)$ | 11, Décharge 1 |

#### 9.2.3.2. *Groupes abéliens et groupes involutifs*

Soit $G$ la théorie dont les axiomes (9.1) ont été donnés plus haut. Toute formule du langage de $G$ peut être $G$-valide, $G$-contingente et $G$-inconsistante. Toute formule $G$-contingente détermine une classe particulière de groupes, et la logique permet de comparer ces classes. On a déjà signalé qu'un groupe est *abélien* s'il vérifie la formule $\forall x \forall y (x.y = y.x)$. Une autre classe intéressante est la classe des groupes *involutifs*, qui vérifient la formule $\forall x (x.x = e)$. Une question se pose naturellement : quel lien y a-t-il entre la classe des groupes abéliens et celle des groupes involutifs? Pour prouver que la seconde est strictement incluse dans la première, il faut, d'abord, déduire le postulat spécifique aux groupes abéliens du postulat spécifique aux groupes involutifs; il faut, ensuite, donner un exemple de groupe abélien non involutif.

Considérons d'abord le premier point; nous en donnons ici une preuve informelle, que le lecteur pourra formaliser à titre d'exercice. Soit $A$ un groupe involutif et soient $x, y$ deux éléments quelconques de $A$. Comme $A$ est involutif, on a $(x.y).(x.y) = e$.

D'autre part, on a aussi $(x.y).(y.x) = (x.(y.y)).x = (x.e).x = x.x = e$. Par transitivité de l'égalité, on a aussi $(x.y).(x.y) = (x.y).(y.x)$; en prémultipliant par l'élément symétrique de $x.y$, on obtient $x.y = y.x$. Ceci établit que tout groupe involutif est abélien. Pour voir que la réciproque est fausse, il suffit de considérer le groupe additif des nombres entiers; il est abélien, puisque $x + y = y + x$ est toujours vrai, mais il n'est pas involutif, puisque $x + x = 0$ n'est vrai que pour $x = 0$.

Ce petit exercice montre comment les formules contingentes donnent lieu à des "raffinements" d'une théorie. Comme nous le verrons dans la suite de ce chapitre, il est intéressant de répéter le processus de raffinement jusqu'à son terme, c'est-à-dire jusqu'à l'obtention d'une théorie pour laquelle il n'existe plus de formules contingentes. Signalons déjà que la théorie des groupes involutifs est de ce type : une formule du langage est soit vraie dans tous les groupes involutifs infinis, soit fausse dans tous ces groupes.

### 9.2.3.3. *Autres présentations de la théorie des groupes abéliens*

Nous avons aussi signalé qu'une même classe de modèles peut être décrite par plusieurs théories. C'est vrai en particulier pour la classe des groupes, et il est commode d'ajouter à la signature de $G$ la constante fonctionnelle $i$ à une place, dont l'interprétation sera la fonction qui associe à tout élément son symétrique. On remplacera alors le postulat

$$\text{A3} : \forall x \exists y \, (x.y = e = y.x)$$

par le postulat

$$\text{A'3} : \forall x \, (x.i(x) = e = i(x).x) \, .$$

Rappelons que ce changement serait impossible si le symétrique d'un élément n'était pas unique. L'adjonction de nouvelles constantes à une théorie permet une formalisation plus aisée de certains résultats, tels les théorèmes bien connus

$$\text{C1} : \forall x \forall y \, (i(x.y) = i(y).i(x)) \, ,$$
$$\text{C2} : \forall x \, (i(i(x)) = x) \, ,$$

dont nous laissons la démonstration au lecteur.

Quand il s'agit de prouver un "métathéorème", c'est-à-dire l'appartenance d'une théorie à une certaine classe de théories, il est parfois plus pratique d'adopter une présentation aussi sobre que possible de la théorie en question. Il est possible, par exemple, de définir une théorie dont la signature ne comporte qu'un élément (une constante fonctionnelle à deux places, notée /), et dont la classe des modèles est isomorphe à la classe des groupes abéliens. Un ensemble adéquat de postulats est

(9.6)
$$\text{B1} : \forall x \forall y \forall z \, \forall u \, (x/(u/(y/z)) = (x/(u/y))/z) \, ,$$
$$\text{B2} : \forall x \forall y \forall z \, (x/(y/z) = z/(y/x)) \, ,$$
$$\text{B3} : \forall x \forall y \, (x/(y/y) = x) \, .$$

*Remarque.* Une conséquence logique intéressante de ces trois postulats est l'énoncé B4 : $\forall x \forall y (x/x = y/y)$. En effet, on a $x/x = (x/(y/y))/x = (y/(y/x))/x = y/(y/(x/x)) = y/y$.

Soit $\mathcal{A}$ la classe des groupes abéliens et soit $\mathcal{B}$ la classe des modèles des axiomes B1, B2 et B3. Il faut montrer que les deux classes sont isomorphes, c'est-à-dire qu'à tout élément $G$ de $\mathcal{A}$ correspond un élément $f(G)$ de $\mathcal{B}$, et réciproquement.

On commence par définir l'isomorphisme $f$. Soit $G = (E, ., e, i)$ un groupe abélien; la structure $f(G)$ sera $(E, /)$, où $a/b =_{def} a.i(b)$. On peut alors vérifier que $f(G)$ vérifie les postulats B1, B2 et B3, dans lesquels l'opération / est exprimée en termes des opérations . et $i$. Il faut donc établir les théorèmes

$$\begin{aligned}
&\text{T1} : \forall x \forall y \forall z \, \forall u \, (x.i(u.i(y.i(z))) = (x.i(u.i(y))).i(z)) , \\
&\text{T2} : \forall x \forall y \forall z \, (x.i(y.i(z)) = z.i(y.i(x))) , \\
&\text{T3} : \forall x \forall y \, (x.i(y.i(y)) = x) .
\end{aligned}$$

Pour le second énoncé, par exemple, on a

$$\begin{aligned}
x.i(y.i(z)) &= x.(i(i(z)).i(y)) , \\
&= x.(z.i(y)) , \\
&= (x.z).i(y) , \\
&= (z.x).i(y) , \\
&= z.(x.i(y)) , \\
&= z.(i(i(x)).i(y)) , \\
&= z/(y/x) .
\end{aligned} \tag{9.7}$$

Le lecteur est invité à vérifier que la structure $f(G)$ vérifie aussi les postulats B1 et B3.

Réciproquement, soit la structure $G' = (E, /)$, un élément de $\mathcal{B}$; on définit la structure $f^{-1}(G') = (E, ., e, i)$, où $a.b =_{def} a/((b/b)/b)$, où $e =_{def} x/x$ et où $i(a) =_{def} (a/a)/a$. Il faut d'abord vérifier que $f^{-1} \circ f$ est la fonction identité dans $\mathcal{A}$ et que $f \circ f^{-1}$ est la fonction identité dans $\mathcal{B}$. (L'opérateur infixé $\circ$ représente la composition[7] de fonctions.)

L'opération binaire, dans la structure $(f^{-1} \circ f)(G)$, associe aux éléments $a, b$ de $G$ l'élément $a/((b/b)/b) = a.i((b.i(b)).i(b))$; cet élément se réduit à $a.i(e.i(b))$, puis à $a.i(i(b))$, c'est-à-dire à $a.b$. Dans la même structure, l'opération unaire associe à l'élément $a$ l'élément $(a/a)/a = (a.i(a)).i(a)$, qui se réduit à $i(a)$. La fonction $f^{-1} \circ f$ est donc bien l'identité dans la classe $\mathcal{A}$. (L'étude de $f \circ f^{-1}$ est laissée au lecteur.)

Il faut enfin vérifier que $f^{-1}(G')$ est bien un groupe abélien. Pour ce faire, il faut montrer que les postulats caractérisant la classe $\mathcal{A}$, dans lesquels on a réécrit les opérations . et $i$ en termes de l'opération /, sont conséquences logiques des postulats caractérisant la classe $\mathcal{B}$. Autrement dit, il faut démontrer les énoncés

[7] Cf. 6.4.10

$(x/((y/y)/y))/((z/z)/z) = x/(((y/((z/z)/z))/(y/((z/z)/z)))/(y/((z/z)/z)))$,
$x/(((x/x)/(x/x))/(x/x)) = x = (x/x)/((x/x)/x)$,
$x/((((x/x)/x)/((x/x)/x))/((x/x)/x)) = x/x = ((x/x)/x)/((x/x)/x)$,
$x/((y/y)/y) = y/((x/x)/x)$.

Voici, à titre d'exemple, une preuve du troisième énoncé :

$$\begin{aligned} x/((((x/x)/x)/((x/x)/x))/((x/x)/x)) &= x/((x/x)/((x/x)/x)), && \text{B4}, \\ &= (x/((x/x)/(x/x)))/x, && \text{B1}, \\ &= x/x, && \text{B3}, \\ &= ((x/x)/x)/((x/x)/x), && \text{B4}. \end{aligned}$$

*Remarque.* Intuitivement, l'unique opération de $G'$ correspond à la division dans $G$.

Le but de cet exercice était de montrer qu'une même classe de structures peut être identifiée à la classe des modèles de plusieurs théories du premier ordre; on verra dans la suite de ce chapitre quel avantage on peut tirer de ce fait. Signalons aussi que certaines classes de structures ne coïncident avec la classe des modèles d'aucune théorie du premier ordre; c'est le cas, notamment, de la classe des groupes finis. Ce fait correspond à une lacune fondamentale de la logique du premier ordre, sur laquelle nous reviendrons dans la suite de ce chapitre.

## 9.3. Théorie de l'ordre

La notion d'ordre intervient dans les domaines les plus variés, non seulement en informatique et en mathématique, mais aussi dans la vie quotidienne. La théorie de l'ordre a pour modèles les ensembles munis d'une relation d'ordre, c'est-à-dire d'une relation réflexive, antisymétrique et transitive. Nous exposons cette théorie non seulement pour sa valeur exemplative, mais aussi pour son intérêt propre. On verra que des notions théoriques importantes s'illustrent facilement dans ce contexte.

Des exemples classiques d'ensembles ordonnés sont les ensembles numériques **N**, **Z**, **Q** et **R**. L'ensemble des suites finies formées à partir d'un alphabet fini est ordonné lexicographiquement. Les ensembles d'objets structurés sont naturellement ordonnés : l'objet $a$ précède l'objet $b$ si $a$ est un constituant (3.3.3) de $b$. L'ensemble des théorèmes déduits d'un système axiomatique donné peut également être ordonné de manière naturelle. Un moyen consiste à fixer une méthode pour engendrer *successivement* tous les théorèmes; on dit alors que le théorème $A$ précède le théorème $B$ si $A$ est engendré avant $B$.

### 9.3.1. *Ordre partiel*

On souhaite définir une théorie $O$ dont les modèles soient exactement les ensembles ordonnés. Le seul objet spécifique du langage sera une constante

prédicative à deux arguments, traditionnellement représentée par le symbole $\leq$ utilisé en notation infixée : $x \leq y$. (Le contexte empêchera généralement toute confusion entre une relation d'ordre générique, et la relation d'ordre numérique également notée $\leq$.) Les trois propriétés caractéristiques des relations d'ordre sont formalisées par les trois postulats classiques :

$$\begin{array}{ll} \forall x\,(x \leq x), & \text{réflexivité}, \\ \forall x \forall y\,((x \leq y \wedge y \leq x) \supset x = y), & \text{antisymétrie}, \\ \forall x \forall y \forall z\,((x \leq y \wedge y \leq z) \supset x \leq z), & \text{transitivité}. \end{array} \tag{9.8}$$

On dispose maintenant d'un moyen rigoureux pour déterminer si un ensemble $K$, muni d'une relation binaire, est ordonné ou non. Considérons la relation "divise" sur les ensembles numériques. Si $E$ est un ensemble numérique, et si $a$ et $b$ sont des éléments de $E$, la relation $a|b$ est vraie si et seulement s'il existe un élément $c$ de $E$ tel que $a.c = b$. La structure $(\mathbf{N}, |)$ est un modèle de $O$, de même que la structure $(\mathbf{N}_0, |)$. Par contre, la structure $(\mathbf{Z}, |)$ n'est pas un modèle de $O$, parce que le postulat d'antisymétrie n'est pas respecté; on a, par exemple, $3|-3$ et $-3|3$, bien que $3 = -3$ soit faux.

La théorie de l'ordre, comme toute théorie, permet de distinguer une sous-classe de la classe des structures mathématiques. On souhaitera souvent affiner ce type de classification dichotomique. Un premier moyen consiste à considérer une formule du langage de la théorie $O$, et à discriminer les ensembles ordonnés, selon qu'ils vérifient ou ne vérifient pas la formule.

### 9.3.2. *Ordre total*

Considérons d'abord la propriété de *totalité*, exprimée par l'énoncé

$$\forall x \forall y (x \leq y \vee y \leq x). \tag{9.9}$$

Cette propriété est vraie dans certains ensembles ordonnés, par exemple dans $(\mathbf{N}, \leq)$ (où $\leq$ dénote ici la relation d'ordre numérique) et fausse dans d'autres, par exemple $(\mathbf{N}, |)$. L'existence de tels énoncés est liée au caractère général de la théorie : il existe une grande variété d'ensembles ordonnés. On dit que la théorie de l'ordre est *incomplète* (7.7.3), parce qu'elle ne rend compte que des propriétés communes à tous les modèles, et non des propriétés relatives à un modèle particulier. Cela signifie que les postulats fournis sont insuffisants pour délimiter une *classe élémentaire* de modèles, c'est-à-dire une classe de modèles dont les éléments ne peuvent pas être discriminés par une formule du langage de la théorie.

L'énoncé de totalité subdivise les ensembles ordonnés en deux catégories : les ordres partiels et les ordres totaux, ou linéaires. Bien souvent, une théorie est développée pour rendre compte d'une certaine réalité; si la théorie est incomplète, la réalité est donc "relativement mal" appréhendée. Dans ce cas, on peut ajouter des postulats, de manière à rendre la théorie plus précise. La

théorie $O'$ de l'ordre total est encore incomplète, comme le montre le postulat de *densité* :

(9.10) $\quad \forall x \forall y (x < y \supset \exists z (x < z \wedge z < y))$.

Ce postulat exprime que, entre deux éléments distincts, il en existe toujours un troisième; il est clair que l'ordre numérique usuel est dense sur l'ensemble des rationnels, mais ne l'est pas sur l'ensemble des entiers. On voit aussi qu'un ensemble muni d'un ordre dense est nécessairement infini, si l'on exclut le cas trivial de l'ensemble vide et des singletons.

La théorie de l'ordre total dense n'est pas encore complète; par exemple, dans $\mathbf{R}$, l'intervalle ouvert $]-1, 1[$ n'admet ni minimum ni maximum, alors que l'intervalle fermé $[-1, 1]$ admet le minimum $-1$ et le maximum 1. Le postulat de *non-bornation* permet de distinguer ces deux cas :

(9.11) $\quad \forall x \exists y \exists z (y < x \wedge x < z)$.

Les modèles de la théorie ainsi obtenue sont les ensembles totalement ordonnés, denses et non bornés, tels les ensembles numériques $\mathbf{Q}$, $\mathbf{R}$ et $]-1, 1[$. Cette théorie, que nous désignerons par $O''$, est *complète*, c'est-à-dire qu'une formule fermée du langage vraie dans un modèle de la théorie est nécessairement un théorème. Autrement dit, toute formule fermée du langage est soit un théorème, soit la négation d'un théorème (complétude syntaxique). Cela revient à dire que les modèles ne peuvent pas être discriminés par des formules du langage.

Les théories complètes sont spécialement intéressantes puisque, dans ce contexte, il suffit de connaître les propriétés d'un modèle pour connaître celles de tous les modèles.

Une technique de preuve de complétude consiste à définir une notion de forme canonique, aussi simple que possible, pour les formules de la théorie. Le concept de forme canonique a déjà été utilisé pour étudier le calcul des propositions, sous le nom de forme normale (conjonctive ou disjonctive). Deux formules sont équivalentes si elles ont même forme canonique; en particulier, une formule est valide si sa forme canonique est $\mathbf{V}$.

La technique de la forme canonique a également été introduite dans le cas du calcul des prédicats (forme prénexe, forme clausale). La réduction sous forme prénexe simplifie l'analyse d'une formule, grâce à la localisation en tête de celle-ci de tous les quantificateurs; l'usage de la forme clausale facilite la mise en œuvre du principe de résolution.

La situation est spécialement favorable dans le cas de la théorie $O''$ de l'ordre total dense et non borné car, dans cette théorie, les quantificateurs peuvent être entièrement éliminés. Autrement dit, dans le cadre de cette théorie, toute formule est équivalente à une formule ouverte, c'est-à-dire à une formule sans quantificateur.

Une relation d'ordre sur l'ensemble fini de variables $E = \{x_1, \ldots, x_n\}$ peut être décrite par la conjonction $\bigwedge_{i,j=1}^{n} l(i,j)$, où $l(i,j)$ est soit $(x_i \leq x_j)$ soit

$\neg(x_i \leq x_j)$. On appelle *description* toute conjonction $O$-consistante de ce type.

*Exemples.* Dans le cas où $n = 1$, il n'existe qu'une description $x_1 \leq x_1$. Dans le cas où $n = 2$, l'ensemble des descriptions comporte quatre éléments qui sont

$$\begin{aligned} &\text{A1}: \quad (x_1 \leq x_2) \wedge (x_2 \leq x_1) \ , \\ &\text{A2}: \quad (x_1 \leq x_2) \wedge \neg(x_2 \leq x_1) \ , \\ &\text{A3}: \quad \neg(x_1 \leq x_2) \wedge (x_2 \leq x_1) \ , \\ &\text{A4}: \quad \neg(x_1 \leq x_2) \wedge \neg(x_2 \leq x_1) \ , \end{aligned}$$

et qui correspondent respectivement aux cas $x_1 = x_2$, $x_1 < x_2$, $x_2 < x_1$ et $x_1$, $x_2$ incomparables. Les trois premières descriptions définissent un ordre total tandis que la quatrième définit un ordre partiel.

*Remarque.* La description A3 se ramène à la description A2 si on échange les rôles des indices 1 et 2. On appelle *permutation* un réarrangement d'un ensemble; plus formellement, une permutation sur $E$ fini est une bijection de $E$ sur lui-même, c'est-à-dire une fonction de $E$ sur $E$ telle que $f(x) = f(y)$ si et seulement si $x = y$. Deux descriptions de $E$ sont *essentiellement distinctes* s'il n'existe pas de permutation de l'ensemble $\{1, \ldots, n\}$ ramenant une description à l'autre. Dans le cas $n = 2$, $E$ admet quatre descriptions et trois descriptions essentiellement distinctes.

Le cas où $n = 3$ est déjà plus complexe. D'une part, quand les trois variables sont comparables (ordre total), les trois variables peuvent être égales (1 cas), deux variables peuvent être égales et la troisième plus grande ou plus petite (6 cas, dont deux essentiellement distincts) ou les trois variables peuvent être différentes (6 cas, non essentiellement distincts). D'autre part, si deux variables sont incomparables, la troisième peut être plus grande, plus petite ou incomparable (9 cas, dont trois essentiellement distincts). Il existe donc 22 descriptions, dont 7 sont essentiellement distinctes.

En toute généralité, la structure de l'ensemble $D(E)$ des descriptions est donnée par le lemme suivant.

**Lemme 9.1.** L'ensemble $D(E)$ est fini. Si $P$ est un sous-ensemble de $D(E)$, on a l'équivalence $(\neg \bigvee(P) \equiv \bigvee(D(E) \setminus P))$[8].

*Démonstration.* Tout d'abord, si $E$ comporte $n$ éléments, une description se compose de $n^2$ littéraux. L'ensemble $D(E)$ comporte au plus $2^{n^2}$ éléments. (Le postulat de réflexivité permet de réduire cette borne supérieure à $2^{n^2-n}$ puisque

[8] Cette formule se lit : "la négation de la disjonction des descriptions membres de $P$ est équivalente à la disjonction des descriptions qui ne sont pas membres de $P$".

le littéral $l(i,i)$ est nécessairement $(x_i \leq x_i)$; la réduction supplémentaire due aux deux autres postulats est plus difficile à évaluer.)
Ensuite, on observe immédiatement que la conjonction de deux descriptions distinctes est inconsistante, car elle implique simultanément un littéral et sa négation. Ceci se généralise facilement et toute conjonction du type

(9.12) $\bigvee(P) \wedge \bigvee(Q)$,

où $P$ et $Q$ sont des sous-ensembles disjoints de $D(E)$, est $O$-inconsistante; ceci implique que la formule

(9.13) $\bigvee(P) \wedge \bigvee(D(E) \setminus P)$

est $O$-inconsistante et que la formule

(9.14) $\bigvee(D(E) \setminus P) \supset \neg \bigvee(P)$

est un théorème. L'implication réciproque se réduit à $\mathbf{V}$, comme le montre le développement suivant :

(9.15) $\neg \bigvee(P) \supset \bigvee(D(E) \setminus P)$,
$\bigvee(P) \vee \bigvee(D(E) \setminus P)$,
$\bigvee D(E)$,
$\mathbf{V}$.

*Exemple.* On se limite ici au cas $n = 2$ et $P = \{\mathrm{A1}, \mathrm{A4}\}$. La formule $\neg \bigvee(P)$ s'écrit et se simplifie comme suit :

$\neg(\mathrm{A1} \vee \mathrm{A4})$,
$\neg(((x_1 \leq x_2) \wedge (x_2 \leq x_1)) \vee (\neg(x_1 \leq x_2) \wedge \neg(x_2 \leq x_1)))$,
$\neg((x_1 \leq x_2) \wedge (x_2 \leq x_1)) \wedge \neg(\neg(x_1 \leq x_2) \wedge \neg(x_2 \leq x_1))$,
$(\neg(x_1 \leq x_2) \vee \neg(x_2 \leq x_1)) \wedge ((x_1 \leq x_2) \vee (x_2 \leq x_1))$,
$(\neg(x_1 \leq x_2) \wedge (x_1 \leq x_2)) \vee (\neg(x_1 \leq x_2) \wedge (x_2 \leq x_1))$
$\quad \vee (\neg(x_2 \leq x_1) \wedge (x_1 \leq x_2)) \vee (\neg(x_2 \leq x_1) \wedge (x_2 \leq x_1))$,
$\mathbf{F} \vee (x_2 < x_1) \vee (x_1 < x_2) \vee \mathbf{F}$,
$(x_2 < x_1) \vee (x_1 < x_2)$.

D'autre part, la formule $\bigvee(D(E) \setminus P)$ devient

$\mathrm{A2} \vee \mathrm{A3}$,
$((x_1 \leq x_2) \wedge \neg(x_2 \leq x_1)) \vee (\neg(x_1 \leq x_2) \wedge (x_2 \leq x_1))$,
$(x_1 < x_2) \vee (x_2 < x_1))$,
$(x_2 < x_1) \vee (x_1 < x_2)$.

Le lemme suivant décrit le pouvoir d'expression des formules ouvertes de la théorie de l'ordre.

**Lemme 9.2.** Dans la théorie de l'ordre, pour toute formule ouverte $A$ dont les variables (libres) appartiennent à l'ensemble $E = \{x_1, \ldots, x_n\}$, il existe un

sous-ensemble $P$ de $D(E)$ pour lequel $A$ est équivalente à $\bigvee(P)$.

*Démonstration.* On réduit d'abord $A$ à la forme normale disjonctive. Il suffit alors de montrer que chaque terme $t$ de la disjonction est équivalent à une disjonction de descriptions sur E. L'ensemble $D(E)$ des descriptions sur $E$ est fini. La disjonction équivalente à $t$ a pour termes les descriptions incluant tous les littéraux de $t$. En effet, soit $\delta$ l'ensemble de ces descriptions. Si $d \in \delta$, alors $(d \supset t)$ est un théorème; on a donc $(\bigvee \delta \supset t)$. On a aussi $(t \supset \bigvee \delta)$, puisque par le lemme 9.1 la formule $(t \wedge \neg \bigvee \delta)$ est équivalente à $(t \wedge \bigvee(D(E) \setminus \delta))$ et est donc inconsistante (chaque description $d \in D(E) \setminus \delta$ contient, par construction, la négation d'un littéral de $t$).

*Exemples.* Considérons, dans le cas $n = 2$, la formule $(x_1 < x_2) \equiv (x_2 < x_1)$. Cette équivalence n'est vraie que si les deux membres sont faux, c'est-à-dire si les deux variables sont égales (description A1) ou incomparables (description A4). Cela apparaît d'ailleurs immédiatement lorsque la formule proposée est réduite à la forme normale disjonctive :

(9.16)
$$\begin{aligned}
&(x_1 < x_2) \equiv (x_2 < x_1)\,,\\
&((x_1 < x_2) \wedge (x_2 < x_1)) \vee (\neg(x_1 < x_2) \wedge \neg(x_2 < x_1))\,,\\
&\mathbf{F} \vee ((\neg(x_1 \leq x_2) \vee x_1 = x_2) \wedge (\neg(x_2 \leq x_1) \vee x_2 = x_1))\,,\\
&(\neg(x_1 \leq x_2) \wedge \neg(x_2 \leq x_1)) \vee ((x_1 \leq x_2) \wedge (x_2 \leq x_1))\,,\\
&\mathrm{A4} \vee \mathrm{A1}\,.
\end{aligned}$$

Considérons aussi, dans le cas $n = 3$, la formule $(x_1 = x_2) \wedge (x_2 \leq x_3)$. Cette formule est équivalente à une forme disjonctive normale à un seul terme, dont les littéraux forment l'ensemble

(9.17) $$A = \{x_1 \leq x_2,\ x_2 \leq x_1,\ x_2 \leq x_3\}\,.$$

Toute description contenant $A$ contiendra aussi $A' = A \cup \{x_1 \leq x_3\}$, car l'ensemble $A \cup \{\neg(x_1 \leq x_3)\}$ est $O$-inconsistant. L'ensemble $A'$ peut a priori être complété de quatre manières, donnant lieu aux ensembles

(9.18)
$$\begin{aligned}
&A' \cup \{x_3 \leq x_1,\ x_3 \leq x_2\}\,,\\
&A' \cup \{\neg(x_3 \leq x_1),\ x_3 \leq x_2\}\,,\\
&A' \cup \{x_3 \leq x_1,\ \neg(x_3 \leq x_2)\}\,,\\
&A' \cup \{\neg(x_3 \leq x_1),\ \neg(x_3 \leq x_2)\}\,.
\end{aligned}$$

Le premier ensemble et le dernier sont des descriptions mais les deux autres n'en sont pas, parce qu'ils sont $O$-inconsistants. La formule initiale est donc la disjonction de deux descriptions.

La notion de description se simplifie dans la théorie de l'ordre total, comme l'exprime le lemme suivant.

**Lemme 9.3.** Toute description $O'$-consistante $d = \bigwedge_{i,j=1}^{n} l(i,j)$ est équivalente à une *description réduite* $d' = \bigwedge_{k=1}^{n-1} l'(d_k, d_{k+1})$, où $(d_1, \ldots, d_n)$ est une permutation de $(1, \ldots, n)$ et où $l'(i,j)$ est $x_i < x_j$ ou $x_i = x_j$.

*Remarque.* En pratique, il est commode de représenter la description $d'$ de l'ensemble $\{x_1, x_2, \ldots, x_n\}$ par une expression du type

(9.19) $\quad x_{d_1} \prec x_{d_2} \prec \ldots \prec x_{d_n}$ ,

où chaque occurrence de $\prec$ est $<$ ou $=$, selon les cas. C'est d'ailleurs un usage fréquent en mathématique d'écrire $(a \prec b \prec c)$ pour $(a \prec b) \wedge (b \prec c)$.

*Démonstration.* La démonstration se fait par induction sur $n$. Si $n$ vaut 1, la seule description consistante est $x_1 \leq x_1$, qui est équivalente à la description réduite $x_1 = x_1$. Soit alors $d =_{def} \bigwedge_{i,j=1}^{n} l(i,j)$ une description. Il existe $k \in \{1, \ldots, n\}$ tel que, pour tout $i \neq k$, le littéral $l(i,k)$ est $(x_i \leq x_k)$. Sinon, il existerait une suite $i_1, \ldots, i_r$ telle que tous les littéraux $\neg(x_{i_1} \leq x_{i_2})$, $\neg(x_{i_2} \leq x_{i_3})$, $\ldots$, $\neg(x_{i_{r-1}} \leq x_{i_r})$, $\neg(x_{i_r} \leq x_{i_1})$ soient vrais, ce qui, dans la théorie $O'$, implique que tous les littéraux $x_{i_2} < x_{i_1}$, $x_{i_3} < x_{i_2}$, $\ldots$, $x_{i_r} < x_{i_{r-1}}$, $x_{i_1} < x_{i_r}$ soient vrais. Cela est impossible, car une conséquence logique de ces littéraux, dans la théorie $O$, est le littéral identiquement faux $x_{i_1} < x_{i_1}$.
La permutation requise sera $(d_1, \ldots, d_n)$, où $d_n = k$ et où $(d_1, \ldots, d_{n-1})$ est la permutation de $(1, \ldots, n-1)$ correspondant à la description initiale dont on a supprimé les littéraux comportant la variable $x_k$. (L'hypothèse d'induction s'applique, car la suppression conduit à une description sur un ensemble de $n-1$ variables.)
Il reste à déterminer les littéraux $l'(d_j, d_{j+1})$. Pour $j < n-1$, ces littéraux seront ceux de la description réduite sur l'ensemble de $n-1$ variables. Le littéral $l'(d_{n-1}, d_n)$ sera $x_{d_{n-1}} = x_k$ si la description initiale comportait le littéral $(x_k \leq x_{d_{n-1}})$ et il sera $x_{d_{n-1}} < x_k$ si la description initiale comportait le littéral $\neg(x_k \leq x_{d_{n-1}})$.

*Exemples.* Considérons l'ensemble $\{x_1, x_2, x_3\}$ et la description

(9.20)
$$\begin{array}{c} (x_1 \leq x_1) \wedge (x_1 \leq x_2) \wedge \quad (x_1 \leq x_3) \wedge \\ \neg(x_2 \leq x_1) \wedge (x_2 \leq x_2) \wedge \neg(x_2 \leq x_3) \wedge \\ (x_3 \leq x_1) \wedge (x_3 \leq x_2) \wedge \quad (x_3 \leq x_3). \end{array}$$

La description réduite s'obtient au moyen de la permutation

$$\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$$

et se représente par l'expression

$$x_1 = x_3 < x_2 .$$

Si on supprime la variable $x_1$, par exemple, la description et la description réduite de l'ensemble résiduel $\{x_2, x_3\}$ sont respectivement

$$(x_2 \leq x_2) \wedge \neg(x_2 \leq x_3) \wedge$$
$$(x_3 \leq x_2) \wedge \ (x_3 \leq x_3)$$

et

$$x_3 < x_2 .$$

Les résultats suivants permettent l'élimination des quantificateurs, dans le cadre de la théorie $O''$.

**Lemme 9.4.** Dans le cadre de la théorie de l'ordre total, dense et non borné, si $A$ est une formule ouverte dont les variables appartiennent à l'ensemble $E$, alors la formule $\exists x_1 A$ est équivalente à une formule ouverte ne comportant pas $x_1$.

*Démonstration.* Comme la fermeture existentielle d'une disjonction est équivalente à la disjonction des fermetures existentielles, il suffit de considérer le cas où $A$ est une description réduite $O''$-consistante (le cas où $A$ est une formule $O''$-inconsistante est trivial). On peut également, au prix d'un renommage des variables, supposer que $A$ est du type $\bigwedge_{k=1}^{n-1} l'(k, k+1)$, où $l'(i,j)$ est soit $x_i < x_j$, soit $x_i = x_j$; on suppose aussi que $x_p$ est le nouveau nom de la variable $x_1$.
On distingue trois cas : $p = 1$, $1 < p < n$ et $p = n$.
Dans le premier cas, la formule se réduit à

$$\exists x_1 \, l'(1,2) \wedge \bigwedge_{k=2}^{n-1} l'(k, k+1) . \tag{9.21}$$

Le terme quantifié peut être $\exists x_1 (x_1 = x_2)$ ou $\exists x_1 (x_1 < x_2)$. La première formule est valide; la seconde est une conséquence de l'axiome de non-bornation. Le terme quantifié peut donc être omis, purement et simplement.
Dans le deuxième cas, la formule $\exists x_p A$ peut se réécrire

$$\bigwedge_{k=1}^{p-2} l'(k, k+1) \wedge \exists x_p \, (l'(p-1, p) \wedge l'(p, p+1)) \wedge \bigwedge_{k=p+1}^{n-1} l'(k, k+1) . \tag{9.22}$$

Si les deux littéraux quantifiés sont des égalités, le terme quantifié est remplacé par $x_{p-1} = x_{p+1}$. Sinon, une $O$-conséquence de la formule $\exists x_p A$ est $x_{p-1} < x_{p+1}$. Vu l'axiome de densité, il existe toujours un $x_p$ tel que les littéraux $l'(p-1, p)$ et $l'(p, p+1)$ soient vrais, ce qui établit que le terme quantifié peut être remplacé par $x_{p-1} < x_{p+1}$.
Le troisième cas est analogue au premier, ce qui achève la démonstration.

*Remarque.* Les variables libres de la formule résultante sont les variables libres de la formule de départ, sauf la variable $x_p$.

*Exemple.* Considérons la formule

$$\mathrm{A} : \exists x_1 (x_2 \leq x_1 \wedge (x_3 < x_1 \supset x_1 < x_2)).$$

En réduisant la matrice (6.4.2)à la forme normale disjonctive, on obtient

$$\exists x_1 ((x_2 \leq x_1 \wedge \neg(x_3 < x_1)) \vee (x_2 \leq x_1 \wedge x_1 < x_2)),$$

ou encore

$$\exists x_1 ((x_2 \leq x_1 \wedge x_1 = x_3) \vee (x_2 \leq x_1 \wedge x_1 < x_3) \vee \mathbf{F}).$$

Cette formule est équivalente à la disjonction des formules

$$\begin{aligned} &\mathrm{A1} : \exists x_1 (x_2 \leq x_1 \wedge x_1 = x_3), \\ &\mathrm{A2} : \exists x_1 (x_2 \leq x_1 \wedge x_1 < x_3), \\ &\mathrm{A3} : \exists x_1 \mathbf{F}. \end{aligned}$$

La formule A3 est identiquement fausse et peut être omise. La matrice de la formule A1 est la disjonction de deux descriptions réduites et il en est de même pour la matrice de la formule A2. La formule initiale A est donc la disjonction des formules suivantes :

$$\begin{aligned} &\mathrm{A11} : \exists x_1 (x_2 = x_1 = x_3), \\ &\mathrm{A12} : \exists x_1 (x_2 < x_1 = x_3), \\ &\mathrm{A21} : \exists x_1 (x_2 = x_1 < x_3), \\ &\mathrm{A22} : \exists x_1 (x_2 < x_1 < x_3). \end{aligned}$$

L'élimination de $x_1$ peut maintenant se faire, en appliquant le procédé indiqué dans la démonstration du lemme 9.4 (deuxième cas). On obtient

$$\begin{aligned} &\mathrm{A11} : (x_2 = x_3), \\ &\mathrm{A12} : (x_2 < x_3), \\ &\mathrm{A21} : (x_2 < x_3), \\ &\mathrm{A22} : (x_2 < x_3). \end{aligned}$$

La formule A de départ est équivalente à la disjonction des formules A11, A12, A21 et A22, c'est-à-dire à la formule $(x_2 \leq x_3)$.

*Remarque.* Dans un but d'illustration, nous avons détaillé le procédé général. La quantification pouvait en fait être éliminée plus simplement, comme le montre le développement suivant :

$$\begin{aligned} &\exists x_1 (x_2 \leq x_1 \wedge (x_3 < x_1 \supset x_1 < x_2)), \\ &\exists x_1 (x_2 \leq x_1 \wedge (\neg(x_3 < x_1) \vee x_1 < x_2)), \\ &\exists x_1 (x_2 \leq x_1 \wedge (x_1 = x_3 \vee x_1 < x_3 \vee x_1 < x_2)), \\ &\exists x_1 ((x_2 \leq x_1 = x_3) \vee (x_2 \leq x_1 < x_3) \vee \mathbf{F}), \\ &x_2 \leq x_3 \vee x_2 < x_3, \\ &x_2 \leq x_3. \end{aligned}$$

**Théorème 9.5.** La théorie de l'ordre total, dense et non borné est complète.

*Démonstration.* Il suffit de montrer que toute formule fermée $B$ est $O''$-valide ou $O''$-inconsistante. Pour ce faire, on met d'abord $B$ en forme prénexe, puis on élimine les quantifications une à une, en commençant par la plus intérieure. On élimine une quantification existentielle en utilisant le lemme 9.4; une quantification universelle est remplacée par une quantification existentielle : $\forall x\, A$ est équivalent à $\neg\exists x \neg A$. A chaque étape, on élimine en même temps une quantification et la variable correspondante; le résultat de ce processus ne peut être qu'une formule sans variable, c'est-à-dire Vou **F**.

On observe que les modèles **R** et $]-1,1[$ de $O''$ sont *isomorphes*, c'est-à-dire qu'il existe entre eux une bijection respectant la relation d'ordre. L'ensemble $]-1,1[$ est l'image de l'ensemble **R** par la bijection

$$(9.23) \quad f : x \to \frac{x}{1+|x|},$$

dont la réciproque est

$$(9.24) \quad f^{-1} : x \to \frac{x}{1-|x|}.$$

On vérifie aisément que cette bijection est un isomorphisme, c'est-à-dire que, pour tout couple de réels $(x,y)$, $x \leq y$ est vrai si et seulement si $f(x) \leq f(y)$ est vrai.

*Remarque.* La notion d'isomorphisme est relative à une théorie donnée; la bijection $f$ respecte la relation d'ordre mais ne respecte pas, par exemple, la multiplication usuelle : $f(x.y)$ n'est pas égal à $f(x).f(y)$.

Une théorie est *catégorique* quand tous ses modèles sont isomorphes. Ce n'est pas le cas de la théorie de l'ordre total, dense et non borné, puisque certains de ses modèles, comme **Q**, sont dénombrables,[9] alors que d'autres, comme **R**, ne le sont pas.

Cantor a cependant établi que tous les modèles dénombrables de la théorie sont isomorphes. Il est commode de démontrer d'abord un lemme.

**Lemme 9.6.** Soient $C$ et $D$ deux ensembles infinis munis d'une relation d'ordre total dense et non borné, et soient $\{c_0,\ldots,c_n\} \subset C$ et $\{d_0,\ldots,d_n\} \subset D$ tels que $c_i \leq c_j \equiv d_i \leq d_j$, pour tous $i,j \leq n$. Pour tout $c_{n+1} \in C$, il existe $d_{n+1} \in D$ tel que $c_i \leq c_j \equiv d_i \leq d_j$, pour tous $i,j \leq n+1$.

*Démonstration.* On distingue quatre cas.

1. S'il existe un $i$ tel que $c_{n+1} = c_i$, on choisit $d_{n+1} = d_i$.
2. Si $c_{n+1} < c_i$ pour tout $i \leq n$, alors l'axiome de non-bornation garantit l'existence d'un élément $d \in D$ tel que $d < d_i$ pour tout $i \leq n$; on choisit

[9] Les expressions "modèle fini" et "modèle dénombrable" sont des abus de langage pour désigner des modèles dont le domaine est fini ou dénombrable; de même, le cardinal d'un modèle est en fait le cardinal du domaine de ce modèle.

$d_{n+1} = d$.

3. De même, si $c_i < c_{n+1}$ pour tout $i \leq n$, il existe $d \in D$ tel que $d_i < d$ pour tout $i \leq n$ et on choisit $d_{n+1} = d$.
4. Sinon, il existe $p, q \in \{0, \ldots, n\}$ tels que $c_p < c_{n+1} < c_q$ et que, pour tout $i \leq n$, soit $c_i \leq c_p$, soit $c_q \leq c_i$. Dans ce cas, vu l'axiome de densité, il existe $d \in D$ tel que $d_p < d < d_q$ et on choisit $d_{n+1} = d$.

**Théorème 9.7.** Tous les modèles dénombrables de la théorie de l'ordre total dense et non borné sont isomorphes.

*Démonstration.* Soient $A$ et $B$ deux ensembles infinis dénombrables, munis d'une relation d'ordre total, dense et non borné. Les ensembles $A$ et $B$ étant infinis dénombrables, il existe des fonctions $a$ et $b$ telles que $A = \{a(i) : i \in \mathbf{N}\}$ et $B = \{b(i) : i \in \mathbf{N}\}$.

On prouve l'existence d'un isomorphisme $f : A \to B$ en le construisant.[10]

On définit inductivement deux suites $(c_n)$ et $(d_n)$, $n \in \mathbf{N}$, telles que, pour tout $n$, les ensembles $C_n = \{c_0, \ldots, c_n\}$ et $D_n = \{d_0, \ldots, d_n\}$ vérifient les propriétés suivantes.

1. $C_n \subset A$ et $D_n \subset B$;
2. $c_i \leq c_j \equiv d_i \leq d_j$, pour tous $i, j \leq n$.

On pose d'abord $C_0 = \{a(0)\}$ et $D_0 = \{b(0)\}$.
Si $C_{2k-1}$ et $D_{2k-1}$ sont définis, on pose $c_{2k} = a(k)$ et on choisit $d_{2k} \in B$ de telle sorte que la propriété 2 soit respectée; le lemme 9.6 garantit qu'un tel choix est toujours possible.
Si $C_{2k}$ et $D_{2k}$ sont définis, on pose $d_{2k+1} = b(k)$ et on choisit $c_{2k} \in A$ de telle sorte que la propriété 2 soit respectée; le lemme 9.6 garantit à nouveau qu'un tel choix est toujours possible.
En alternant de la sorte les choix dans $A$ et dans $B$, on réalise les inclusions $\{a(0), \ldots, a(n)\} \subset C_{2n}$ et $\{b(0), \ldots, b(n)\} \subset D_{2n+1}$, pour tout $n$. On en déduit les égalités $A = \bigcup_n C_n = \{c_i : i \in \mathbf{N}\}$ et $B = \bigcup_n D_n = \{d_i : i \in \mathbf{N}\}$. On définit alors un isomorphisme $f : A \to B$ en posant $f(c_i) = d_i$, pour tout $i \in \mathbf{N}$.

Si $\alpha$ est un cardinal, on dit qu'une théorie est *$\alpha$-catégorique* si tous ses modèles de cardinal $\alpha$ sont isomorphes. Le théorème de Cantor qui vient d'être introduit exprime donc que la théorie de l'ordre total dense et non borné est $\omega$-catégorique. Le symbole $\omega$ représente ici le cardinal de l'ensemble $\mathbf{N}$; une théorie est donc $\omega$-catégorique si tous ses modèles infinis dénombrables sont isomorphes.

[10] La méthode de construction utilisée est dite "de long en large" ou "en zig-zag" (*"back and forth"* en anglais). Elle a été introduite par Cantor et est d'usage fréquent en théorie des modèles.

L'adjonction du postulat de non-bornation à la théorie de l'ordre total dense a donné lieu à une théorie complète. Les théories complètes obtenues par adjonction de postulats supplémentaires à la théorie de l'ordre total dense sont les suivantes :

- la théorie de l'ordre total dense non borné,
  dont l'axiome spécifique est : $\forall x \exists y \exists z (y < x \wedge x < z)$ ;
- la théorie de l'ordre total dense borné à gauche seulement,
  dont l'axiome spécifique est : $\exists g \forall x (g \leq x) \wedge \forall x \exists z (x < z)$ ;
- la théorie de l'ordre total dense borné à droite seulement,
  dont l'axiome spécifique est : $\exists d \forall x (x \leq d) \wedge \forall x \exists y (y < x)$ ;
- la théorie de l'ordre total dense borné à gauche et à droite,
  dont l'axiome spécifique est : $\exists g \exists d \forall x (g \leq x \wedge x \leq d)$.

La démonstration donnée pour le premier cas s'adapte facilement aux trois autres; on peut également prouver (moins facilement) que la liste donnée ci-dessus est exhaustive.

Le fait que des structures aussi différentes que $\mathbf{Q}$ et $\mathbf{R}$ ne puissent être distinguées l'une de l'autre dans le cadre de la théorie de l'ordre ne traduit pas, en soi, une insuffisance de la logique formelle. En effet, la distinction entre les réels et les rationnels est possible dans d'autres théories, contenant par exemple la multiplication en plus de la relation d'ordre; la formule

(9.25) $\quad \forall x (0 \leq x \supset \exists y (y.y = x))$

est vraie dans $\mathbf{R}$ mais fausse dans $\mathbf{Q}$. Elle reste cependant vraie dans certains ensembles dénombrables, comme celui des nombres algébriques.

On peut quand même considérer qu'une certaine lacune existe, dans la mesure où les relations d'ordre sur $\mathbf{Q}$ et $\mathbf{R}$ sont discernables. Considérons par exemple l'ensemble des nombres positifs dont le carré est compris entre 2 et 3. Dans $\mathbf{R}$, cet ensemble admet un plus petit élément $\sqrt{2}$ et un plus grand élément $\sqrt{3}$, alors que dans $\mathbf{Q}$, ces éléments extrémaux n'existent pas. Ce résultat ne contredit pas les précédents : il montre seulement que la notion d'intervalle, quoique directement liée au concept de relation d'ordre, ne peut pas s'exprimer dans le cadre de la logique classique. Cette notion s'exprime par contre facilement dans des logiques d'ordre supérieur.

### 9.3.3. *Bon ordre*

La notion d'*ordre total* est introduite pour rendre compte des cas où deux éléments d'un ensemble sont toujours comparables : l'un est le premier et l'autre est le second. Le postulat spécifique de la théorie de l'ordre total, c'est-à-dire la formule $\forall x \forall y (x \leq y \vee y \leq x)$, traduit cette propriété. Ceci se généralise immédiatement : tout sous-ensemble comportant $n$ éléments donne lieu à une suite telle que chaque terme est supérieur à ses prédécesseurs et inférieur à ses successeurs; la notion de description réduite introduite au paragraphe précédent

met cette propriété en évidence. Un corollaire immédiat de cette propriété est l'existence, pour tout sous-ensemble fini d'un ensemble ordonné, d'un premier élément, appelé le *minimum* du sous-ensemble. En effet, pour tout entier $n$ supérieur à 2, on peut déduire dans la théorie de l'ordre total la formule

$$(9.26)\quad \forall x_1 \cdots \forall x_n \quad \exists y \quad ((y = x_1 \vee \cdots \vee y = x_n) \wedge y \leq x_1 \wedge \cdots \wedge y \leq x_n).$$

Il est naturel de s'intéresser alors aux ensembles *bien ordonnés* dont, par définition, *tous* les sous-ensembles non vides admettent un minimum. Un *bon ordre* est nécessairement un ordre total, puisque le postulat spécifique de la théorie de l'ordre total exprime que toute paire d'éléments admet un minimum. La réciproque est vraie dans le cas des ensembles finis (on montre aisément par récurrence sur $n$ que si toute paire admet un minimum, alors tout $n$-uplet admet un minimum) mais est fausse dans le cas général. Les ensembles $\mathbf{Q}$ et $\mathbf{R}$ sont totalement ordonnés sans être bien ordonnés. Dans les deux cas, l'intervalle ouvert $]0, 1[$ n'admet ni minimum ni maximum. En fait, la propriété d'existence d'un minimum pour tout sous-ensemble non vide est incompatible avec la propriété de densité. Si $E$ est un ensemble totalement ordonné, dense et non borné, et si $x \in E$, alors l'ensemble des éléments plus grands que $x$ n'admet pas de minimum.

Ce résultat négatif donne lieu à la question suivante : existe-t-il un ensemble de formules qui, réuni aux postulats de la théorie de l'ordre total, fournisse un ensemble de postulats dont la classe des modèles serait exactement la classe des ensembles bien ordonnés ? La réponse est à nouveau négative, comme nous le verrons plus loin (théorème 9.14).

### 9.3.4. *Le principe d'induction*

Nous avons formulé et utilisé à plusieurs reprises (3.4.9, 3.4.16, 5.4.6, 8.4.2) dans les chapitres précédents des variantes du *principe d'induction*, qui permet d'établir des propriétés communes à tous les éléments de certains ensembles ordonnés, comme l'ensemble des nombres naturels, ou l'ensemble des formules du calcul des propositions (cf. 3.3.2). Des liens étroits existent entre la notion d'ensemble ordonné et le principe d'induction, ce qui justifie une étude spécifique de ce principe dans le contexte de la théorie de l'ordre.

On a beaucoup discuté le statut de ce principe. Sa particularisation à l'ensemble des entiers naturels prend le nom de *principe de récurrence* et peut s'énoncer comme suit :

> Si un énoncé est vrai pour le nombre 0,
> et si de l'énoncé pour un nombre naturel quelconque $n$,

on peut déduire l'énoncé pour le nombre $n + 1$,
alors l'énoncé est vrai pour tous les nombres naturels.

Utilisé et explicitement reconnu par Pascal en 1654 et par Fermat en 1659 (Quine, 1940, p. 243), le principe de récurrence a été longtemps tenu pour un axiome mathématique irréductible à la logique. Une affirmation explicite de cette position se trouve dans *la Science et l'Hypothèse*, (1902), où Poincaré fait observer que le principe du raisonnement par récurrence diffère des principes de logique en cela qu'on peut réduire n'importe quel principe de logique au principe de contradiction,[11] mais qu'on ne peut en faire autant pour le principe du raisonnement par récurrence, ce dernier enveloppant une infinité de raisonnements. Le principe du raisonnement par récurrence, déployé, génère une suite infinie :

L'énoncé est vrai du nombre 0
Or s'il est vrai de 0, il est vrai de 1
Donc il est vrai de 1
Or s'il est vrai de 1, il est vrai de 2
Donc il est vrai de 2 et ainsi de suite
(Poincaré, 1902, 1968, p. 39).

Poincaré en conclut que la règle d'induction mathématique, inaccessible à la démonstration analytique et à l'expérience, est le véritable type du jugement synthétique *a priori*.[12] Poincaré entend par jugement synthétique a priori une proposition qui accroît notre savoir (contrairement au jugement analytique, qui explicite un savoir implicite) et qui est cependant antérieur à l'expérience.

On peut facilement ramener le principe du raisonnement par récurrence au principe de contradiction si on dispose des notions de nombre naturel et d'ensemble des nombres naturels mais ces notions n'appartiennent pas au domaine classique de la logique. Dans *Logique et Connaissance Scientifique*, Piaget s'abstient de faire appel au concept aujourd'hui abandonné de jugement synthétique a priori[13], mais reprend en partie à son compte la thèse de

[11] Par exemple, le Modus Ponens "$P$ et si $P$ alors $Q$ donc $Q$" peut se lire "affirmer que $P$ et que si $P$ alors $Q$ et nier $Q$, c'est transgresser le principe de contradiction".

[12] C'est Kant qui a introduit la notion de jugement synthétique a priori, dans *la Critique de la Raison Pure*. Un jugement synthétique a priori est à la fois synthétique et a priori.

> "Ou le prédicat B appartient au sujet A comme quelque chose qui est contenu implicitement dans ce concept A, ou B est entièrement en dehors du concept A, quoiqu'il soit, à la vérité, en connexion avec lui. Dans le premier cas, je nomme le jugement *analytique*, dans l'autre *synthétique*." (Kant, 1781, 1950, p. 37)

> "Si [...] on trouve une proposition dont la pensée implique la *nécessité*, on a un jugement *a priori*." (Kant, ibid, p. 33)

[13] Si le concept est abandonné, il ne s'ensuit pas que la position de Kant ne puisse pas être défendue sous une forme aménagée. Voir "Kant vindicated" dans J. Hintikka, 1973 et J. P. Dubucs *Grazer Philosophische Studien* 1988

Poincaré :

> Par contre, Poincaré avait certainement raison en soutenant que [ ... ] le raisonnement par récurrence est bien plus "fort" qu'une déduction non appuyée sur la série infinie des entiers (Piaget, 1967, p. 72).

On doit concéder à Poincaré qu'il n'est pas possible de faire du principe d'induction mathématique un énoncé analytique au sens étroit où "analytique" signifie "réductible au principe de contradiction". Cependant, le problème se pose en des termes différents si on prend "analytique" au sens (défini par Frege) de "réductible, *par des définitions*, à une loi de logique".

Comme nous le verrons au chapitre suivant (volume II), un cadre adéquat pour donner et exploiter des définitions formalisées est la théorie des ensembles; il est donc intéressant d'étudier le principe de récurrence dans ce cadre, ce que nous ferons.

Il est néanmoins possible de formaliser le principe d'induction dans toute théorie permettant de définir une relation d'ordre ou, plus généralement, une relation antisymétrique $\rho$. Les symboles $\preceq$ et $\prec$ seront utilisés pour désigner respectivement les versions réflexive et antiréflexive de cette relation. Formellement, on pose

$$(9.27) \qquad \begin{aligned} (x \prec y) &=_{def} ((x \rho y) \wedge (x \neq y)), \\ (x \preceq y) &=_{def} ((x \rho y) \vee (x = y)). \end{aligned}$$

*Remarques.* Les symboles $\leq$ et $<$ sont réservés aux relations antisymétriques et transitives.

Un seul des symboles $\prec$ et $\preceq$ est nécessaire, l'autre pouvant être considéré comme une abréviation. On a en effet

$$(9.28) \qquad \begin{aligned} (x \prec y) &\equiv ((x \preceq y) \wedge (x \neq y)), \\ (x \preceq y) &\equiv ((x \prec y) \vee (x = y)). \end{aligned}$$

Le principe d'induction est souvent utilisé dans le cadre de théories plus riches que la théorie de l'ordre, dans lesquelles la relation $\prec$ (ou $\preceq$) est définie au moyen d'autres symboles. Par exemple, dans le cadre de la théorie des nombres naturels, on peut poser $(x \prec y) \ =_{def} \ (x + 1 = y)$, ou encore $(x \leq y) =_{def} \exists z\,(x + z = y)$.

Nous nous plaçons dans une théorie $T$ permettant la définition d'une relation antisymétrique, dont nous utiliserons la version antiréflexive, notée $\prec$. Le *principe d'induction* est la règle d'inférence suivante :

$$(9.29) \qquad \frac{\forall x\,((\forall y\,((y \prec x) \supset R(y))) \supset R(x))}{\forall x\,R(x)}$$

Il s'agit en fait d'un schéma de règle d'inférence, dans lequel $R(x)$ est toute formule de la théorie ne comportant pas d'autre variable libre que $x$. L'idée consiste à prouver la propriété $R$ pour un certain $x$ à partir de la même propriété supposée déjà établie pour les éléments $y$ précédant $x$. Intuitivement, cette règle

ne sera valable que si le domaine de la théorie ne contient pas de suite infinie décroissante, c'est-à-dire de suite $(x_1, \ldots, x_n \ldots)$ telle que $x_{n+1} \prec x_n$ pour tout entier naturel $n$. Une suite infinie décroissante donnerait lieu à une "régression à l'infini".

Il n'est pas possible de formaliser la notion de suite dans le cadre des théories du premier ordre, sauf en faisant usage de techniques spéciales (comme dans le cadre de la théorie des ensembles). Par contre, la notion de sous-ensemble se formalise très bien. Toute formule $S(x)$ n'admettant aucune variable libre autre que $x$ définit un sous-ensemble du domaine d'interprétation : celui des individus $a \in \mathcal{D}$ tels que $S(x)$ est vrai quand l'interprétation de $x$ est $a$.

*Remarque.* On effectue ici une incursion déguisée dans la théorie des ensembles, mais cette incursion (licite : on ne sort pas de la théorie du premier ordre $T$) est très limitée et non dangereuse. Tout d'abord, seuls des sous-ensembles bien particuliers du domaine d'interprétation sont accessibles. En effet, si ce domaine est infini (et le cas d'un domaine fini n'est guère intéressant), la classe des sous-ensembles définissables sera au plus dénombrable, alors que la classe de tous les sous-ensembles du domaine est non dénombrable. D'autre part, la consistance de la théorie $T$ garantit l'absence de paradoxes, tel le paradoxe russellien de la classe des classes qui ne se contiennent pas elles-mêmes. Même si la consistance de la théorie $T$ n'est pas établie, le principe de séparation de Zermelo ("*Aussonderung Axiom*") s'applique ici : le domaine d'interprétation étant un ensemble, c'est-à-dire une classe non paradoxale, toutes les classes définissables seront également non paradoxales.

La *condition de minimalité*

$$(9.30) \quad \exists x\, S(x) \supset \exists z\, (S(z) \wedge \forall y\, (S(y) \supset \neg(y \prec z)))$$

exprime que le sous-ensemble du domaine caractérisé par $S$, s'il n'est pas vide, admet un *élément minimal* $z$, c'est-à-dire un élément qui n'est précédé par aucun autre élément du sous-ensemble. Une suite décroissante dont les éléments forment l'ensemble caractérisé par $S$ ne peut contenir un élément minimal que comme dernier élément; elle ne peut donc pas être infinie.

Dans la suite, la condition de minimalité est notée MIN, l'antécédent du principe d'induction est noté API et son conséquent est noté CPI.

**Théorème 9.8.** Si la condition de minimalité est satisfaite pour toute formule $S(x)$, alors le principe d'induction est une règle d'inférence valide.

*Démonstration.* Un schéma de preuve par l'absurde est donné ci-dessous.

| | | |
|---|---|---|
| 1. | $\exists x \neg R(x)$ | $\neg$ CPI |
| 2. | $\exists x \neg R(x) \supset \exists z\, (\neg R(z) \wedge \forall y\, (\neg R(y) \supset \neg(y \prec z)))$ | MIN $(\neg R)$ |
| 3. | $\exists z\, (\neg R(z) \wedge \forall y\, (\neg R(y) \supset \neg(y \prec z)))$ | 1, 2 |
| 4. | $\exists z\, (\neg R(z) \wedge \forall y\, (y \prec z \supset R(y)))$ | 3 |
| 5. | $\exists z\, (\neg R(z) \wedge R(z))$ | 4, API |

Dans le cadre de la théorie des nombres naturels, on peut poser

(9.31) $(x \prec y) =_{def} (x + 1 = y)$.

La condition de minimalité devient

(9.32) $\exists x\, S(x) \supset \exists z\,(S(z) \wedge \forall y(S(y) \supset (y + 1 \neq z)))$.

Cette condition est toujours satisfaite; un élément minimal adéquat est le minimum (au sens habituel) de l'ensemble non vide $\{r : S(r)\}$. Le principe d'induction est donc valide et devient

(9.33) $$\frac{\forall x\,((\forall y\,((y + 1 = x) \supset R(y))) \supset R(x))}{\forall x\, R(x)}.$$

On sait que, dans $\mathbf{N}$, tout nombre $x$ a un et un seul prédécesseur (noté $x - 1$), sauf 0 qui n'en a pas. L'antécédent de la règle se réécrit successivement comme suit. (Le lecteur est invité à restituer les étapes manquantes du développement.)

1. $\forall x\,((\forall y\,((y + 1 = x) \supset R(y))) \supset R(x))$
2. $\forall x\,(x = 0 \supset ((\forall y\,((y + 1 = x) \supset R(y))) \supset R(x))) \wedge$
   $\forall x\,(x \neq 0 \supset ((\forall y\,((y + 1 = x) \supset R(y))) \supset R(x)))$
3. $((\forall y\,((y + 1 = 0) \supset R(y))) \supset R(0)) \wedge \forall x\,((x \neq 0 \supset (R(x - 1))) \supset R(x)))$
4. $((\forall y\,(\mathbf{F} \supset R(y))) \supset R(0)) \wedge \forall u\,((\mathbf{V} \supset (R(u + 1 - 1))) \supset R(u + 1)))$
5. $((\forall y\,\mathbf{V}) \supset R(0)) \wedge \forall u\,(\mathbf{V} \supset (R(u + 1 - 1) \supset R(u + 1)))$
6. $R(0) \wedge \forall x\,(R(x) \supset R(x + 1))$

et le principe d'induction devient

(9.34) $$\frac{R(0) \wedge \forall x\,(R(x) \supset R(x + 1))}{\forall x\, R(x)},$$

ce qui est le principe de récurrence habituel.

De même, le choix $(x \prec y) =_{def} (x < y)$, pour lequel la condition de minimalité est également satisfaite, permet de retrouver la règle classique de l'induction complète.

### 9.3.5. *Applications*

Au terme de cette section relative à la théorie de l'ordre, le lecteur pourrait estimer que l'approche logique est à la fois laborieuse et aléatoire. Laborieuse, parce que les résultats positifs, comme la complétude de la théorie de l'ordre total, dense et non borné, s'obtiennent par le biais de démonstrations difficiles; aléatoire, parce que d'autres résultats sont négatifs, notamment le fait que les bons ordres ne forment la classe de modèles d'aucune théorie du premier ordre.

Nous voudrions montrer ici qu'une telle vision est exagérément pessimiste, et que les obstacles mentionnés peuvent être surmontés.

Il est clair, tout d'abord, que la notion même d'ordre est complexe, bien qu'elle se résume en trois postulats simples. (Cette concision est à mettre à l'actif de la logique formelle, et non à son passif!) Notre premier exercice sera donc de rendre plus intuitive cette notion, en fournissant pour elle une représentation standard. Comme bien souvent en mathématique, une telle représentation appartiendra à la théorie élémentaire des ensembles, dite aussi théorie naïve, telle qu'elle est utilisée dans la majorité des ouvrages mathématiques contemporains. Le problème est plus précisément celui-ci : construire, dans le cadre de la théorie naïve des ensembles, une relation d'ordre $R$, telle que, pour tout ensemble ordonné $(E, \leq)$, il existe un ensemble ordonné $(\mathcal{E}, R)$ isomorphe à $(E, \leq)$. Si on note $f$ cet isomorphisme, on doit avoir

(9.35) $\quad \mathcal{E} = \{f(x) : x \in E\}$ [14],
$\forall x, y \in E\,(x = y \equiv f(x) = f(y))$,
$\forall x, y \in E\,(x \leq y \equiv f(x)\, R\, f(y))$.

Un premier point est que la notion d'ordre intervient très tôt dans le développement de la théorie naïve des ensembles. La relation de base de cette théorie, l'appartenance, est antisymétrique mais non transitive. Toutefois, on introduit immédiatement la relation d'inclusion, selon le schéma

(9.36) $\quad (a \subset b) =_{def} \forall x\,(x \in a \supset x \in b)$ ;

cette dernière relation est une relation d'ordre, puisqu'elle est, par construction, réflexive, antisymétrique et transitive.

Un second point, plus utile que le premier, est que toute relation d'ordre $\leq$ sur un ensemble $E$ est fidèlement représentée par la relation d'inclusion sur une famille d'ensembles

(9.37) $\quad \mathcal{E} =_{def} \{f(x) : x \in E\}$ où $f(x) =_{def} \{y : (y \in E \wedge y \leq x)\}$.

On associe à chaque élément $x$ de $E$ l'ensemble $f(x)$ des éléments de $E$ plus petits que lui. L'intérêt de cette représentation est que la notion d'ordre, a priori très diverse, peut être exhaustivement étudiée à partir de la notion intuitive et fondamentale d'inclusion. La théorie des ensembles fera l'objet du chapitre suivant (volume II), mais les concepts élémentaires suffisent à évoquer des questions intéressantes. Tout d'abord, la propriété critique d'un bon ordre est l'existence d'un minimum pour tout sous-ensemble non vide. Si la relation d'ordre est l'inclusion, ce minimum admet une expression immédiate : le minimum d'une famille d'ensembles $\mathcal{F}$ est l'intersection $\bigcap \mathcal{F}$ de ses éléments. Une famille d'ensembles est donc bien ordonnée par inclusion si elle contient toute intersection de ses éléments. L'assimilation d'une relation d'ordre arbitraire à la relation d'inclusion sur une certaine famille d'ensembles se révèle féconde à plusieurs titres, et nous aurons l'occasion d'y revenir plus en détail dans le volume II.

[14] Cette égalité se lit : l'ensemble $\mathcal{E}$ est l'ensemble des valeurs que prend la fonction $f$ quand les valeurs de l'argument sont les éléments de l'ensemble $E$.

Prouver la complétude d'une théorie $T$ ou, plus généralement, déterminer exhaustivement les adjonctions de postulats qui donnent lieu à des théories complètes contenant $T$, n'est pas chose facile. Néanmoins, la technique d'élimination des quantificateurs qui a été utilisée plus haut est d'application assez générale[15]. Les difficultés de son application ne sont d'ailleurs pas liées à la logique elle-même, mais à la théorie particulière étudiée. Rappelons le sens du mot "complet" appliqué à une théorie : une théorie (formalisée) est complète si pour toute formule fermée, il y a une preuve soit de celle-ci, soit de sa négation, et si toute formule vraie dans un modèle de la théorie est vraie dans tous les modèles (ce sont là les versions syntaxique et sémantique de la complétude syntaxique (cf. 7.7.3).

Enfin, il convient de noter qu'en logique formelle, bien souvent, un résultat négatif est à la base de résultats positifs. En ce qui concerne la notion de bon ordre, le fait qu'elle sorte du cadre de la logique du premier ordre est un indice de sa richesse. Cela apparaîtra plus clairement lors de l'étude de la théorie des ensembles, où le concept d'ensemble bien ordonné a une importance cruciale.

Plus concrètement, le résultat négatif qui vient d'être rappelé — à savoir le fait que les bons ordres ne forment la classe des modèles d'aucune théorie du premier ordre — suggère l'exercice suivant : définir une classe élémentaire d'ensembles ordonnés, aussi restreinte que possible mais incluant tous les ensembles bien ordonnés.

Puisque la condition d'existence d'un minimum pour tout sous-ensemble non vide est trop forte, on peut chercher à l'affaiblir et à ne l'exiger que pour des sous-ensembles particuliers. Si l'existence du minimum n'est exigée que des sous-ensembles finis, on obtient la classe élémentaire des ordres totaux. On peut aussi envisager le cas où l'ensemble des *majorants* $E_x =_{def} \{y : x < y\}$ d'un élément $x$ admet un minimum. Ce minimum est appelé le *successeur* de $x$. La notion de successeur de $x$ est du premier ordre, car elle est caractérisée par la formule suivante :

$$(9.38) \quad succ(x) = y =_{def} (x < y \wedge \forall z (x < z \supset y \leq z)).$$

Il est donc possible de définir la théorie de l'ordre avec successeur (unique). Cette théorie n'est pas complète, notamment parce qu'elle n'implique pas la totalité, c'est-à-dire la formule $\forall x \forall y (x \leq y \vee y \leq x)$. Un troisième essai est la théorie de l'ordre total borné à gauche avec successeur unique. Elle n'est pas complète non plus, car un élément peut admettre 0 ou 1 prédécesseur ($a$ est le prédécesseur de $b$ si $b$ est le successeur de $a$).

A ce stade, il est difficile d'aller plus loin car il existe des ensembles bien ordonnés admettant des éléments (autres que l'élément initial) sans prédécesseur.

[15] Pour être précis, signalons que cette technique a été quelque peu généralisée; le point important est qu'un petit nombre de méthodes permet l'étude systématique des théories, au point de vue de la complétude.

(On verra au chapitre suivant que c'est le cas de tous les ordinaux infinis, à l'exception du premier.)

Ce développement suggère néanmoins la définition de la théorie de l'ordre avec successeur et prédécesseur uniques, avec élément initial (le lecteur est invité à écrire les postulats). Comme dans le cas de l'ordre total dense, on peut vérifier que cette théorie est complète. Elle n'est évidemment pas catégorique (cf. 7.6.2), mais la classe de ses modèles peut se décrire comme suit. A tout ensemble totalement ordonné $A$, on associe le modèle $\mathcal{M}_A$ comme suit. Le domaine de $\mathcal{M}_A$ est $\mathbf{N} \cup (A \times \mathbf{Z})$, et la relation d'ordre $R$ est donnée par trois règles simples : l'ordre dans $\mathbf{N}$ est l'ordre habituel, l'ordre dans $(A \times \mathbf{Z})$ est l'ordre lexicographique induit par l'ordre donné sur $A$ et l'ordre habituel sur $\mathbf{Z}$, et enfin tous les éléments de $\mathbf{N}$ précèdent tous les éléments de $(A \times \mathbf{Z})$. Formellement, si on se donne les éléments $m, n \in \mathbf{N}$, $x, y \in \mathbf{Z}$ et $a, b \in A$, ces trois règles s'écrivent :

$$
\begin{aligned}
(9.39) \qquad & m \, R \, n \iff m \leq n \,, \\
& (a, x) \, R \, (b, y) \iff (a \leq b \vee (a = b \wedge x \leq y)) \,, \\
& m \, R \, (a, x) \,.
\end{aligned}
$$

Ce dernier résultat — à savoir la complétude de la théorie de l'ordre avec successeur et prédécesseur uniques et avec élément initial — est positif et intéressant, même s'il n'est pas directement relié à la notion de bon ordre.

## 9.4. Théorie des modèles

La brève étude que nous avons faite de la théorie de l'ordre illustre bien comment la logique est utilisée pour modéliser une certaine réalité. On appelle *métathéorie des modèles* (par abus de langage, théorie des modèles), la branche de la logique consacrée aux liens existant entre les structures mathématiques (algèbres universelles) et les théories (ensembles de formules). Le problème central consiste à déterminer si, oui ou non, telle structure mathématique est un modèle de telle théorie. A partir de ce type de question, on peut en concevoir beaucoup d'autres; nous en donnons un échantillon ci-dessous.

- Quel est l'ensemble des modèles de la théorie $T$ ?
- Etant donné un ensemble $\mathcal{E}$ de structures, existe-t-il une théorie dont l'ensemble des modèles soit exactement $\mathcal{E}$ ?
- Quel lien y a-t-il entre les théories complètes et les théories catégoriques ?

### 9.4.1. *Deux résultats essentiels*

La théorie des modèles repose en grande partie sur deux résultats très importants, le théorème de compacité et le théorème de Löwenheim-Skolem. Nous

introduisons d'abord deux définitions et un lemme, tous deux valables à la fois pour le calcul des propositions et le calcul des prédicats.

Un ensemble de formules est *finiment consistant*, ou *f-consistant*, si tous ses sous-ensembles finis sont consistants. Un ensemble f-consistant est *maximal* si, pour toute formule $\phi$ appartenant au langage considéré, l'une des formules $\phi$ ou $\neg\phi$ appartient à l'ensemble.

*Remarques.* Tout ensemble consistant est f-consistant. L'ensemble des formules vraies dans un modèle donné est f-consistant maximal.

**Théorème 9.9.** Tout ensemble f-consistant est consistant.

*Démonstration.* Soit $E$ un ensemble f-consistant. D'après le théorème de complétude[16], si $E$ n'admet pas de modèle, alors on peut prouver une contradiction à partir des éléments de $E$. La preuve $(A_1, \ldots, A_n \vdash B \wedge \neg B)$ est nécessairement finie et ne fait intervenir qu'un nombre fini d'éléments de $E$. Ces éléments forment un sous-ensemble fini inconsistant de $E$ (à savoir $\{A_1, \ldots, A_n\}$), ce qui est contraire à l'hypothèse.

Le théorème 9.9, appelé *théorème de compacité*, est de nature purement sémantique et ne dépend pas de la notion de système axiomatique. Il est donc possible de le démontrer sans faire référence au théorème de complétude. La démonstration se fait en deux étapes. On montre d'abord que tout ensemble f-consistant est inclus dans un ensemble f-consistant maximal, et ensuite que tout ensemble f-consistant maximal admet un modèle.

**Lemme 9.10.** Un ensemble f-consistant peut toujours être inclus dans un ensemble f-consistant maximal.

*Démonstration.* Soit $E$ un ensemble de formules dont tout sous-ensemble fini admet un modèle, et soit $\mathcal{L}$ l'ensemble des formules construites avec les symboles apparaissant dans les éléments de $E$. On suppose ici que $\mathcal{L}$ est un ensemble dénombrable, ce qui permet d'écrire $\mathcal{L} = \{A_i : i \in \mathbf{N}\}$. On définit par récurrence une famille $\{F_i : i \in \mathbf{N}\}$ d'ensembles de formules, au moyen des deux règles suivantes :

- $F_0 = E$;

[16] dont l'une des formulation est : tout ensemble de formules du calcul des prédicats dont on ne peut déduire une inconsistance a un modèle (cf. 8.4.2). De cette formulation découlent comme *corollaires* :

1. si $E \models A$ alors $E \vdash A$ et
2. si $\models A$ alors $\vdash A$

qui sont parfois présentés comme des formulations du théorème de complétude plutôt que comme corollaires.

- Si tous les sous-ensembles finis de $F_i \cup \{A_i\}$ sont consistants, alors $F_{i+1} = F_i \cup \{A_i\}$, sinon, $F_{i+1} = F_i \cup \{\neg A_i\}$.

On prouve aisément par induction que, pour tout $i \in \mathbf{N}$, l'ensemble $F_i$ est f-consistant. Le seul point non évident est que si $F_i \cup \{A_i\}$ admet un sous-ensemble fini inconsistant $F_1 \cup \{A_i\}$, alors tout sous-ensemble fini $F_2$ de $F_i \cup \{\neg A_i\}$ est consistant. On raisonne comme suit. L'ensemble $F_1 \cup F_2 \setminus \{\neg A_i\}$ est inclus dans $F_i$ et, vu l'hypothèse de récurrence, est consistant. Dans tout modèle $\mathcal{M}$ de cet ensemble, la formule $A_i$ est fausse (sinon $\mathcal{M}$ serait un modèle de $F_1 \cup \{A_i\}$) et $\mathcal{M}$ est donc un modèle de $F_1 \cup F_2$, ce qui implique la consistance de $F_2$. L'ensemble $F_i$ est f-consistant, quel que soit $i$. En conséquence, l'ensemble $F = \bigcup_{i \in \mathbf{N}} F_i$ est également f-consistant (tout sous-ensemble fini de $F$ est nécessairement inclus dans un $F_i$); de plus, il est maximal par construction.

*Remarque.* Dans cette démonstration, on a utilisé les procédés de définition et de raisonnement par récurrence, dans lesquels l'ensemble bien ordonné **N** joue un rôle essentiel. Le défaut de cette approche est d'être limitée aux collections dénombrables. Il est cependant possible d'utiliser, au lieu de **N**, n'importe quel ensemble bien ordonné (dénombrable ou non). On parle alors d'*induction ordinale*, ou d'*induction transfinie*. Une introduction claire et rigoureuse de ces notions ne peut se faire que dans le cadre de la théorie des ensembles (volume II) mais signalons déjà que ces techniques permettent de supprimer l'hypothèse de dénombrabilité du langage. Notons également que l'extension des procédés de récurrence au cas non dénombrable est cruciale en mathématique, et renforce l'importance des notions d'ordre et de bon ordre.

Voici maintenant un second lemme.

**Lemme 9.11.** Tout ensemble f-consistant maximal est consistant.

On considère séparément le cas du calcul des propositions et celui du calcul des prédicats. Le premier cas est très simple car, visiblement, les interprétations du calcul des propositions et les ensembles f-consistants maximaux sont en correspondance bijective. A chaque interprétation $\mathcal{I}$, on associe l'ensemble $g(\mathcal{I})$ des formules vraies dans $\mathcal{I}$; inversement, à chaque ensemble f-consistant maximal $E$, on associe l'interprétation $g^{-1}(E)$ définie comme suit : si $p$ est une proposition, alors $p$ est vraie pour $g^{-1}(E)$ si $p \in E$, et $p$ est fausse pour $g^{-1}(E)$ si $p \notin E$, c'est-à-dire si $\neg p \in E$. On déduit immédiatement que, dans le calcul des propositions, tout ensemble f-consistant maximal admet un et un seul modèle.

On peut ramener le cas du calcul des prédicats à celui du calcul des propositions en utilisant le théorème de Herbrand (8.3). On peut aussi, dans le cas du calcul des prédicats avec égalité, procéder de manière plus directe et plus intuitive. Une formule existentielle du type $\exists x\, \phi(x)$ est vraie s'il existe un objet $c$ tel que $\phi(c)$ est vraie. Si le langage comporte une constante individuelle $c$

pour désigner cet objet, alors $\exists x\, \phi(x)$ est équivalent à $\phi(c)$. On dit que $c$ est une *constante-témoin* pour la formule $\exists x\, \phi(x)$ (cf. déf. 6.8). Pour peu qu'il existe une constante-témoin pour chaque formule de ce type, on dispose d'un moyen pour éliminer les quantificateurs et pour se ramener au cas du calcul des propositions (ce que fait déjà la théorie de Herbrand).

Dans ce contexte, il est assez simple d'associer un modèle à un ensemble $E$ f-consistant maximal. Le domaine $C$ du modèle sera le quotient[17] de l'ensemble des constantes-témoins par la relation d'équivalence $\approx$ définie comme suit : on a $c_1 \approx c_2$ si et seulement si la formule $c_1 = c_2$ appartient à l'ensemble $E$. On convient de désigner une classe d'équivalence par n'importe lequel de ses représentants.

*Remarque.* Le fait que $\approx$ est une relation d'équivalence est une conséquence immédiate du fait que $E$ est f-consistant maximal. La formule $c_1 = c_1$ est valide; elle appartient donc à $E$, et la relation $\approx$ est réflexive. Si la formule $c_1 = c_2$ appartient à $E$, la formule $c_2 = c_1$ qui en découle appartiendra aussi à $E$; la relation $\approx$ est donc symétrique. Enfin, si les formules $c_1 = c_2$ et $c_2 = c_3$ appartiennent à $E$, alors leur conséquence logique $c_1 = c_3$ appartient à $E$, ce qui montre la transitivité de la relation $\approx$.

Il reste à associer, à toute constante fonctionnelle $f$ d'arité $n$ du langage, une fonction $f'$ de $C^n$ dans $C$ et, à toute constante prédicative $P$ d'arité $m$, une fonction caractéristique $P'$ de $C^m$ dans $\{\mathbf{V}, \mathbf{F}\}$. Ceci se fait simplement : l'élément $f'(c_1, \ldots, c_n)$ est $c$ si et seulement si la formule $f(c_1, \ldots, c_n) = c$ appartient à l'ensemble $E$. (Puisque $\exists x\, f(c_1, \ldots, c_n) = x$ est valide, un élément $c$ adéquat existe toujours.) De même, la fonction propositionnelle $P'(c_1, \ldots, c_m)$ aura la valeur $\mathbf{V}$ si et seulement si la formule $P(c_1, \ldots, c_m)$ fait partie de l'ensemble $E$.

Une démonstration plus formelle de ce résultat apparaît par exemple dans (Chang et Keisler, 1973, p. 61 et sq.), où l'on montre aussi que l'hypothèse d'existence des constantes-témoins n'est pas une réelle restriction.

En combinant les lemmes 9.10 et 9.11, on obtient le théorème de compacité, énoncé ci-dessus (théorème 9.9).

*Remarque.* La dénomination usuelle du théorème de compacité vient de ce qu'il exprime en fait qu'un certain espace topologique (qui n'a pas été introduit ici) est compact; nous reviendrons sur ce point dans le paragraphe consacré aux applications.

Le théorème de compacité est à la source de nombreux résultats négatifs,

[17] Quand on divise un ensemble au moyen d'une relation d'équivalence (symétrique, réflexive et transitive), on obtient une partition, c'est-à-dire un ensemble de classes d'équivalence, ensemble dont les membres sont disjoints et non vides. On appelle ensemble quotient cette partition. Sont membres de la même classe d'équivalence les membres de l'ensemble initial qui sont reliés entre eux par la relation d'équivalence précitée.

exprimant les limitations de la logique du premier ordre. Le premier résultat concerne l'inaptitude de la logique du premier ordre à discriminer les modèles finis des modèles infinis; plus formellement, on a le théorème suivant.

**Théorème 9.12.** Soit $T$ une théorie. Si $T$ admet des modèles finis de taille arbitrairement grande, alors $T$ admet au moins un modèle infini.

*Démonstration.* Soit $C = \{c_i : i \in \mathbf{N}\}$ un ensemble de constantes n'appartenant pas au langage de $T$. Dans le langage augmenté de l'ensemble $C$, on considère la théorie $T' =_{def} T \cup \{c_i \neq c_j : (i,j \in \mathbf{N} \wedge i < j)\}$. Cette théorie est f-consistante car pour tout sous-ensemble fini $A$ de $T'$, il existe un entier $n$ tel que les formules de $T'$ ne comportent pas de constantes d'indice supérieur à $n$; un modèle $\mathcal{M} = (\mathcal{D}, \mathcal{I})$ de $T$ donne lieu immédiatement à un modèle $\mathcal{M}' = (\mathcal{D}, \mathcal{I}')$ de $A$, dès que $\#\mathcal{D} > n$. En conséquence, $T'$ admet un modèle, nécessairement infini, à partir duquel on obtient un modèle infini de $T$.

On démontre d'une manière analogue le *théorème de Lōwenheim et Skolem*, qui exprime l'inaptitude de la logique du premier ordre à discriminer les modèles infinis de cardinaux différents. Voici l'énoncé de ce théorème.

**Théorème 9.13.** Soit $T$ une théorie. Si $T$ admet un modèle infini, alors $T$ admet un modèle de cardinal $\alpha$, pour tout cardinal infini $\alpha$.

Une conséquence de ce théorème est que toute théorie admettant un modèle infini admet à la fois des modèles dénombrables et des modèles non dénombrables. La logique du premier ordre ne se prête donc pas parfaitement à l'étude des ensemble numériques usuels, comme $\mathbf{N}$ et $\mathbf{R}$, puisqu'elle ne peut pas distinguer directement le dénombrable du non dénombrable. On verra, au chapitre suivant, comment la théorie des ensembles permet de contourner cette difficulté.

*Remarque.* Certains auteurs considèrent le cas des théories dont le langage $\mathcal{L}$ n'est pas dénombrable. Dans ce cas, l'existence d'un modèle infini implique l'existence d'un modèle infini de cardinal $\alpha$, pour tout $\alpha \geq |\mathcal{L}|$.

**Théorème 9.14.** La notion de bon ordre n'est pas axiomatisable dans la théorie de l'ordre.

*Démonstration.* Soit $T_b$ une théorie de signature $\{\leq\}$, dont tous les ensembles bien ordonnés sont des modèles. On doit prouver que $T_b$ admet aussi un modèle non bien ordonné, c'est-à-dire un modèle comportant une suite décroissante infinie.

Soit $C = \{c_i : i \in \mathbf{N}\}$ un ensemble de constantes et soit $A$ l'ensemble des formules $\{c_{i+1} < c_i : i \in \mathbf{N}\}$. On considère alors la théorie $T_b \cup A$, de signature $\{\leq\} \cup C$. Il est clair que tout modèle de cette théorie sera un modèle de $T_b$ (il suffira de ne plus interpréter les $c_i$), mais ne sera pas un bon ordre. Par le *théorème de compacité*, un tel modèle existe si tout sous-ensemble fini $E \subset T_b \cup A$ admet un modèle, ce que nous allons vérifier.

Si $E$ est fini, il existe un entier $r$ tel que tous les $c_i$ intervenant dans $E$ vérifient $i < r$. Un modèle de $E$ est l'ensemble N des entiers naturels, muni de l'ordre habituel, où $c_i$ est interprété en $r - i$ si $i < r$ et en 0 sinon.

### 9.4.2. *Relations entre théories*

La relation la plus simple que l'on puisse définir entre deux théories $T$ et $T'$ de même langage $\mathcal{L}$ est celle d'*inclusion*. La théorie $T$ est *incluse* dans la théorie $T'$ si tous les théorèmes de $T$ sont des théorèmes de $T'$. On dit aussi que $T'$ est une *extension* de $T$, mais il est préférable d'éviter ce terme qui recevra une autre acception dans le paragraphe suivant.

Nous avons vu que la théorie de l'ordre est incluse dans la théorie de l'ordre total, elle-même incluse dans la théorie de l'ordre total dense et non-borné.

On peut généraliser la notion d'inclusion comme suit. Une théorie $T$ de langage $\mathcal{L}$ est incluse dans une théorie $T'$ de langage $\mathcal{L}'$ si toute formule de $\mathcal{L}$ est une formule de $\mathcal{L}'$ et si tout théorème de $T$ est un théorème de $T'$. Cette notion est d'usage fréquent en algèbre, où, par exemple, la théorie des demi-groupes, la théorie des monoïdes, la théorie des groupes, la théorie des anneaux, la théorie des corps et la théorie des champs (corps commutatifs) forment une suite dont chaque élément est inclus dans l'élément suivant.

Ces notions, pour utiles qu'elles soient, rendent mal compte des questions réelles qui surviennent lors de la comparaison de deux théories. Par exemple, nous avons montré que la théorie des groupes pouvait être présentée dans différents langages. De telles variantes ne sont pas directement comparables par inclusion, mais les classes correspondantes de modèles sont isomorphes. Il est donc intéressant de comparer des théories dont les langages ne sont pas directement comparables. Cela peut se faire au moyen de la notion de *représentabilité*. Nous n'aborderons cette notion que par le biais d'exemples issus de la théorie des nombres, puisque c'est en ce domaine que la notion de représentabilité s'est avérée la plus féconde. La théorie des nombres (naturels) est construite au moyen d'une collection relativement importante, et d'ailleurs non clairement fixée, de constantes prédicatives et fonctionnelles. Il y a bien sûr les opérations arithmétiques fondamentales, la relation d'ordre numérique, la notion de nombre premier et celles de nombres premiers entre eux, les notions de plus grand commun diviseur et de plus petit commun multiple, la notion d'égalité modulo $k$, etc. Ces notions ne sont pas indépendantes et, depuis Peano, l'usage s'est répandu de considérer l'addition et la multiplication comme les notions fondamentales, dont dérivent toutes les autres. Dans certains cas, la dérivation est facile à effectuer; voici quelques exemples :

$$
\begin{array}{lll}
(9.40) & x = 0 & =_{def}\ x = x + x\,, \\
& x \leq y & =_{def}\ \exists z\,(x + z = y)\,, \\
& x < y & =_{def}\ x \leq y \wedge \neg(x = y)\,, \\
& x = 1 & =_{def}\ (0 < x \wedge \forall y\,(y < x \equiv y = 0))\,, \\
& x|y & =_{def}\ \exists z\,(x.z = y)\,, \\
& gcd(x, y) = z & =_{def}\ (z|x \wedge z|y \wedge \forall u\,((u|x \wedge u|y) \supset u|z))\,, \\
& \ldots & =_{def}\ \ldots
\end{array}
$$

Par contre, d'autres notions importantes, comme l'exponentiation et la factorielle, s'introduisent par récurrence:

$$
\begin{array}{lll}
(9.41) & x^0 & =_{def}\ 1\,, \\
& x^{y+1} & =_{def}\ x.x^y \\
& 0! & =_{def}\ 1\,, \\
& (x + 1)! & =_{def}\ (x + 1).x!\,.
\end{array}
$$

Le problème est que le mécanisme de définition par récurrence est très puissant, et que l'introduction de nouvelles constantes fonctionnelles et prédicatives par cette technique modifie radicalement les propriétés des théories qui en font l'objet.

Un exemple célèbre est celui de l'arithmétique additive, due à Presburger. Le langage de la théorie de Presburger est construit sur l'addition et toutes les constantes prédicatives et fonctionnelles qui se définissent (sans récurrence!) à partir de l'addition. Il est possible de définir sur ce langage une théorie complète, dont les théorèmes sont exactement les formules du langage vraies dans la structure $(\mathbf{N}, +)$. Par contre, si le mécanisme de récurrence est autorisé, on peut définir la multiplication et, à partir de celle-ci, les autres notions usuelles de l'arithmétique introduites plus haut; on a en effet $x.0 = 0$ et $x.(y+1) = x.y + x$.

Gödel a montré que l'adjonction de la multiplication à la théorie de Presburger anéantissait toutes les propriétés intéressantes de cette dernière. Tout d'abord, l'ensemble des formules appartenant au langage étendu et vraies sur la structure classique $(\mathbf{N}, +, .)$ n'est pas récursif ni même récursivement énumérable. Il n'est donc pas possible de déterminer algorithmiquement si une formule du langage étendu est vraie ou non dans $\mathbf{N}$, alors que c'est possible pour une formule du langage de Presburger.

Bien plus, même la théorie classique dite de Peano, bien que ne permettant pas d'obtenir toutes les vérités de l'arithmétique, reste incomplète et même indécidable. Par contre, Gödel a montré que, dans cette théorie, le mécanisme de définition par récurrence devient inutile. En particulier, il est possible de définir l'exponentiation et la factorielle sans récurrence (mais les définitions sont lourdes et font appels à des résultats difficiles de l'arithmétique).

En dehors de son intérêt intrinsèque, la théorie des nombres est importante pour l'informaticien. Elle intervient directement dans certains domaines de l'informatique comme la cryptographie, mais aussi indirectement chaque fois

qu'une structure infinie dénombrable est étudiée. En effet, de telles structures peuvent être rendues isomorphes à **N** par *codage*. La question se pose alors de savoir si on est dans un monde de type Presburger (complet et décidable) ou de type Peano (incomplet et indécidable).

La notion de représentabilité reste très utile dans ce contexte; on peut chercher à "renforcer" la théorie de Presburger (tout en conservant la décidabilité et la complétude), ou à "affaiblir" celle de Peano (tout en conservant le pouvoir d'expression élevé de la théorie de Peano). Cette dernière démarche a donné lieu à de nombreux résultats. Par exemple, on n'affaiblit pas la théorie de Peano en remplaçant l'opération binaire d'addition par l'opération unaire successeur car, comme l'a observé Tarski, on a $S(x.z).S(y.z) = S(z.z.S(x.y))$ si et seulement si $x + y = z$ ou $z = 0$.

### 9.4.3. *Relations entre structures*

Lors de l'étude d'une théorie, on est souvent amené à considérer diverses structures se rapportant au langage de la théorie; il est parfois opportun de modifier ce langage (par adjonction de nouvelles constantes, notamment) et aussi la théorie, par adjonction de nouveaux postulats. Ces modifications se répercutent naturellement sur les structures considérées.

Ces structures peuvent être comparées par divers moyens que nous allons introduire.

#### 9.4.3.1. *Structures définies sur un langage commun*

Deux structures $\mathcal{M} = (\mathcal{D}, \mathcal{I})$ et $\mathcal{M}' = (\mathcal{D}', \mathcal{I}')$ sont *isomorphes* si elles se rapportent au même langage $\mathcal{L}$ et s'il existe entre elles un $\mathcal{L}$-*isomorphisme*, c'est-à-dire une bijection $f : \mathcal{D} \mapsto \mathcal{D}'$ vérifiant les conditions suivantes :

1. Pour toute constante prédicative $P$ à $n$ places et pour tous $d_1, \ldots, d_n \in \mathcal{D}$, l'expression $\mathcal{I}(P)(d_1, \ldots, d_n)$ est vraie si et seulement si l'expression $\mathcal{I}'(P)(f(d_1), \ldots, f(d_n))$ est vraie.
2. Pour toute constante fonctionnelle $g$ à $m$ places et pour tous $d_0, \ldots, d_m \in \mathcal{D}$, l'expression $\mathcal{I}(g)(d_1, \ldots, d_n)$ vaut $d_0$ si et seulement si l'expression $\mathcal{I}'(g)(f(d_1), \ldots, f(d_n))$ vaut $f(d_0)$.

La notion de $\mathcal{L}$-isomorphisme généralise les notions courantes d'isomorphismes. Soit par exemple $\mathcal{L}$ le langage engendré par la signature $\{u, f\}$ où $u$ est une constante individuelle et $f$ une constante fonctionnelle à deux places. Deux structures $\mathcal{L}$-isomorphes sont $(\mathbf{R}, 0, +)$ et $(\mathbf{R}_0^+, 1, .)$; un $\mathcal{L}$-isomorphisme est la fonction exponentielle. En effet, cette fonction est une bijection de l'ensemble des réels sur l'ensemble des réels strictement positifs; de plus, on a $e^0 = 1$ et $e^{x+y} = e^x.e^y$.

Deux structures isomorphes sont véritablement indiscernables[18] : elles ne diffèrent que par les symboles utilisés pour désigner leur constituants.

Il est clair que deux structures $\mathcal{L}$-isomorphes déterminent la même partition des formules fermées de $\mathcal{L}$ en formules vraies et en formules fausses. La réciproque n'est pas toujours vraie, et on dit que deux structures $\mathcal{M}$ et $\mathcal{M}'$ sont *$\mathcal{L}$-équivalentes* si elles rendent vraies les mêmes formules fermées de $\mathcal{L}$; ceci se note $\mathcal{M} \equiv \mathcal{M}'$. La distinction entre isomorphisme de modèles et équivalence de modèles souligne les limites relatives de la logique du premier ordre. On a vu plus haut que les structures $(\mathbf{Q}, \leq)$ et $(\mathbf{R}, \leq)$ sont équivalentes sans être isomorphes. On a cependant le théorème suivant.

**Théorème 9.15.** Deux structures équivalentes telles que tout élément du domaine soit l'interprétation d'une constante du langage sont isomorphes.

*Démonstration.* Soient $\mathcal{M} = (\mathcal{D}, \mathcal{I})$ et $\mathcal{M}' = (\mathcal{D}', \mathcal{I}')$ deux structures équivalentes de langage $\mathcal{L}$. On suppose que la signature de ce langage comporte un ensemble $C$ de constantes individuelles tel que l'on ait $\mathcal{D} = \{\mathcal{I}(c) \ : \ c \in C\}$ et $\mathcal{D}' = \{\mathcal{I}'(c) \ : \ c \in C\}$. Le $\mathcal{L}$-isomorphisme $f$ de $\mathcal{M}$ dans $\mathcal{M}'$ est donné par la condition suivante: $f(d) = d'$ si et seulement s'il existe une constante $c \in C$ telle que $\mathcal{I}(c) = d$ et $\mathcal{I}'(c) = d'$. Supposons par exemple que $P$ soit une constante prédicative à $n$ places appartenant à la signature du langage. Il faut montrer que, pour tous $d_1, \ldots, d_n \in \mathcal{D}$, $\mathcal{I}(P)(d_1, \ldots, d_n)$ est vrai si et seulement si $\mathcal{I}'(P)(f(d_1), \ldots, f(d_n))$ est vrai. Vu l'hypothèse, ces deux expressions peuvent se réécrire respectivement sous la forme $\mathcal{I}(P(c_1, \ldots, c_n))$ et $\mathcal{I}'(P(c_1, \ldots, c_n))$, pour des constantes $c_1, \ldots, c_n \in C$ adéquates. Or ces deux expressions sont équivalentes, vu que $P(c_1, \ldots, c_n)$ est une formule du langage et que $\mathcal{M}$ et $\mathcal{M}'$ sont équivalentes. Un raisonnement analogue s'applique aux constantes fonctionnelles du langage.

Dans le même ordre d'idée, on a aussi le théorème suivant.

**Théorème 9.16.** Deux structures finies équivalentes sont isomorphes.

*Démonstration.* Soient $\mathcal{A}$ et $\mathcal{B}$ deux structures finies équivalentes. On prouvera seulement qu'elles ont même cardinal (condition nécessaire mais non suffisante d'isomorphisme). Soit $E_n$ la formule $\exists x_1 \cdots \exists x_n \forall x (x = x_1 \vee \cdots \vee x = x_n)$.

---

[18] On pourrait cependant soutenir que tout en étant *indiscernables*, ces deux structures sont cependant *différentes*. On pourrait invoquer le fait que $\mathbf{R}_0^+$ est un sous-ensemble propre de $\mathbf{R}$, c'est-à-dire qu'il est contenu dans $\mathbf{R}$ mais non identique à $\mathbf{R}$. (Deux ensembles ne sont identiques que s'ils ont tous leurs membres en commun.) *Cette différence entre $\mathbf{R}$ et $\mathbf{R}^+$ n'est pas uniquement terminologique.* Si on admet cette différence, on doit rejeter le principe leibnizien de *l'identité des indiscernables* qui s'énonce ainsi :

$$\forall x \forall y (\forall \Phi (\Phi x \equiv \Phi y) \supset x = y).$$

La formule $F_n =_{def} (E_n \wedge \neg E_{n-1})$ caractérise les structures dont le cardinal est $n$. Le cardinal commun $r$ de $\mathcal{A}$ et $\mathcal{B}$ est l'indice de l'unique formule $F_r$ vérifiée par $\mathcal{A}$ et $\mathcal{B}$.

Une *sous-structure* de $\mathcal{M} = (\mathcal{D}, \mathcal{I})$ est une structure $\mathcal{M}' = (\mathcal{D}', \mathcal{I}')$, où $\mathcal{D}'$ est un sous-ensemble de $\mathcal{D}$ et où l'interprétation $\mathcal{I}'$ est la restriction à $\mathcal{D}'$ de l'interprétation $\mathcal{I}$. Observons que tout sous-ensemble $\mathcal{D}'$ de $\mathcal{D}$ ne donne pas nécessairement lieu à une sous-structure, sauf si le langage ne comporte pas de constante fonctionnelle. En effet, si $f$ est une constante fonctionnelle à $n$ places, il faut que le codomaine de la restriction à $\mathcal{D}'^n$ de $\mathcal{I}(f)$ soit inclus dans $\mathcal{D}'$; dans le cas contraire, la fonction $\mathcal{I}'(f)$ ne serait pas partout définie. De plus, une sous-structure $\mathcal{M}'$ d'un modèle $\mathcal{M}$ d'une théorie $T$ n'est pas toujours un modèle de $T$ (si c'est le cas, $\mathcal{M}'$ est un *sous-modèle* de $\mathcal{M}$ pour la théorie $T$).

Si $\mathcal{M}'$ est une sous-structure de $\mathcal{M}$, ce que l'on note $\mathcal{M}' \subset \mathcal{M}$, on dit que $\mathcal{M}$ est une *extension*[19] de $\mathcal{M}'$. Si de plus, pour toute formule $\psi$ à $n$ variables libres, on a $\mathcal{M} \models \psi(d_1, \ldots, d_n)$ si et seulement si $\mathcal{M}' \models \psi(d_1, \ldots, d_n)$, quels que soient $d_1, \ldots, d_n \in \mathcal{D}'$, on dit que $\mathcal{M}'$ est une *sous-structure élémentaire* de $\mathcal{M}$, ou que $\mathcal{M}$ est une *extension élémentaire* de $\mathcal{M}'$, ce que l'on note $\mathcal{M}' \prec \mathcal{M}$. Si on a $\mathcal{M}' \prec \mathcal{M}$, on a aussi $\mathcal{M}' \equiv \mathcal{M}$, mais la réciproque est fausse.

La notion de sous-structure élémentaire est très forte. Etant donné une théorie $T$ admettant le modèle $\mathcal{M}$, une sous-structure $\mathcal{M}'$ de $\mathcal{M}$ ne sera pas nécessairement un modèle de $T$. Dans le même ordre d'idée, A. Robinson a introduit la définition suivante. Une théorie $T$ est dite *complète pour les modèles* quand tout sous-modèle d'un modèle $\mathcal{M}$ de $T$ est un sous-modèle élémentaire de $\mathcal{M}$. Une caractérisation simple existe pour cette notion.

**Théorème 9.17.** Une théorie $T$ est complète pour les modèles si et seulement si, dans cette théorie, toute formule purement existentielle[20] est équivalente à une formule purement universelle.

Le critère suivant détermine quand une sous-structure est élémentaire.

**Théorème 9.18.** Soit $\mathcal{M}' = (\mathcal{D}', \mathcal{I}')$ une sous-structure de $\mathcal{M} = (\mathcal{D}, \mathcal{I})$. On a $\mathcal{M}' \prec \mathcal{M}$ si et seulement si, pour toute formule $\phi(x_0, \ldots, x_n)$, pour tout $d_i \in \mathcal{D}$ et pour tous $d'_1, \ldots, d'_n \in \mathcal{D}'$ tels que $\phi(d'_0, d'_1, \ldots, d'_{i-1}, d_i, d'_{i+1}, \ldots, d'_n)$ soit vrai dans $\mathcal{M}$, il existe un $d'_i \in \mathcal{D}'$ tel que $\phi(d'_0, d'_1, \ldots, d'_{i-1}, d'_i, d'_{i+1}, \ldots, d'_n)$ soit également vrai dans $\mathcal{M}$.

*Démonstration.* Pour simplifier l'écriture, on considère seulement le cas $i = 0$. Supposons d'abord que la sous-structure $\mathcal{M}'$ de $\mathcal{M}$ est élémentaire. Par

[19] Il faut distinguer soigneusement le mot "extension" appliqué à des théories du même mot appliqué à des structures.

[20] C'est-à-dire du type $\exists x_1 \cdots \exists x_n \phi$, où $\phi$ ne contient pas de quantificateurs.

application de la définition à la formule $\psi =_{def} \exists x_0\, \phi(x_0, \ldots, x_n)$, on obtient

(9.42) $\quad \mathcal{M} \models \exists x_0\, \phi(x_0, d'_1, \ldots, d'_n) \iff \mathcal{M}' \models \exists x_0\, \phi(x_0, d'_1, \ldots, d'_n)$.

Si les deux membres sont vrais, alors il existe un $d'_0$ tel que $\phi(d'_0, d'_1, \ldots, d'_n)$ soit vrai dans $\mathcal{M}'$ et donc dans $\mathcal{M}$. Si au contraire les deux membres sont faux, alors il n'existe aucun $d_0 \in \mathcal{D}$ tel que $\phi(d_0, d'_1, \ldots, d'_n)$ soit vrai dans $\mathcal{M}$. Dans les deux cas, la condition énoncée dans le théorème est satisfaite.

Inversement, supposons que $\mathcal{M}$ et sa sous-structure $\mathcal{M}'$ vérifient la condition de l'énoncé. On doit montrer que, pour toute formule $\psi(x_0, \ldots, x_n)$, $\psi(d'_0, \ldots, d'_n)$ est vrai dans $\mathcal{M}$ si et seulement si $\psi(d'_0, \ldots, d'_n)$ est vrai dans $\mathcal{M}'$, quels que soient $d'_0, \ldots, d'_n \in \mathcal{D}'$. Cela se démontre par induction sur la structure de la formule $\psi$. Si $\psi$ est atomique, la valeur commune des deux expressions interprétées est $\mathcal{I}'(\psi)(d'_0, \ldots, d'_n)$. Si $\psi$ est une combinaison booléenne $C(\psi_0, \ldots, \psi_m)$ de formules $\psi_0, \ldots, \psi_m$ dont les interprétations en $\mathcal{M}$ et en $\mathcal{M}'$ coïncident, la valeur commune de $\psi(d'_0, \ldots, d'_n)$ en $\mathcal{M}$ et en $\mathcal{M}'$ sera $C(\mathcal{I}'(\psi_0), \ldots, \mathcal{I}'(\psi_m))$. Il reste à montrer que si la coïncidence a lieu pour la formule $\psi$, elle a encore lieu pour la formule $\exists x_i\, \psi$ (à nouveau, on supposera $i = 0$). Deux cas sont possibles. D'une part, s'il existe $d_0 \in \mathcal{D}$ tel que $\psi(d_0, d'_1, \ldots, d'_n)$ soit vrai dans $\mathcal{M}$, alors, vu la condition, il existe $d'_0 \in \mathcal{D}'$ tel que $\psi(d'_0, d'_1, \ldots, d'_n)$ soit vrai dans $\mathcal{M}'$; dans ce cas, $\exists x_0\, \psi(x_0, d'_1, \ldots, d'_n)$ sera vrai dans $\mathcal{M}$ et dans $\mathcal{M}'$. D'autre part, si un tel $d_0 \in \mathcal{D}$ n'existe pas, la formule $\exists x_0\, \psi(x_0, d'_1, \ldots, d'_n)$ sera fausse dans $\mathcal{M}$ comme dans $\mathcal{M}'$.

### 9.4.3.2. *Structures définies sur des langages distincts*

Si la signature du langage $\mathcal{L}$ est incluse dans celle du langage $\mathcal{L}'$, on dit que $\mathcal{L}'$ est une *expansion* de $\mathcal{L}$; c'est une *expansion simple* si seules des constantes individuelles ont été ajoutées à la signature. Etant donné une expansion $\mathcal{L}'$ du langage $\mathcal{L}$, on peut toujours enrichir une structure $\mathcal{M}$ de langage $\mathcal{L}$ en une structure $\mathcal{M}'$ de langage $\mathcal{L}'$; il suffit d'ajouter une interprétation pour les nouvelles constantes. L'*expansion* $\mathcal{M}'$ ainsi obtenue n'est pas unique, mais la *réduction* $\mathcal{M}$ de $\mathcal{M}'$ est visiblement unique.

*Remarque.* Une structure $\mathcal{M}$ de langage $\mathcal{L}$ peut être enrichie de deux manières: par expansion et par extension. On construit une expansion d'une structure en augmentant la signature du langage; on construit une extension d'une structure en augmentant le domaine de la structure.

Un type d'expansion fréquemment utilisé est le suivant. Soit $\mathcal{M} = (\mathcal{D}, \mathcal{I})$ une structure de langage $\mathcal{L}$. L'expansion $\mathcal{L}_\mathcal{D}$ de $\mathcal{L}$ s'obtient par adjonction à la signature de $\mathcal{L}$ d'une constante individuelle $c_d$ pour tout $d \in \mathcal{D}$. L'*expansion* $\mathcal{M}_\mathcal{D}$, souvent notée $(\mathcal{M}, d)_{d \in \mathcal{D}}$, est la structure de langage $\mathcal{L}_D$ obtenue à partir de $\mathcal{M}$ en fixant $\mathcal{I}(c_d) = d$.

On peut utiliser cette notation pour reformuler la définition du concept

d'extension élémentaire. L'extension $\mathcal{M} = (\mathcal{D}, \mathcal{I})$ de $\mathcal{M}' = (\mathcal{D}', \mathcal{I}')$ est élémentaire si et seulement si les expansions $\mathcal{M}_{\mathcal{D}'}$ et $\mathcal{M}'_{\mathcal{D}'}$ sont équivalentes.

A titre de premier exemple, observons que, avec les notations classiques, $(\mathbf{Z}, +)$ est une extension de $(\mathbf{N}, +)$, mais que $(\mathbf{Z}, -)$ n'est pas une extension de $(\mathbf{N}, -)$; en fait, $(\mathbf{N}, -)$ n'est pas une structure, puisque la différence de deux entiers naturels n'est pas toujours un entier naturel. Une expansion de $(\mathbf{Z}, +)$ est $(\mathbf{Z}, +, *)$.

Toutes ces notions permettent d'identifier et de formaliser les liens entre une théorie et ses modèles. Donnons sans démonstration deux exemples d'un tel lien. Voici d'abord une définition : une théorie est *inductive* si tous ses axiomes sont de la forme

$$(9.43) \quad \forall x_1 \cdots \forall x_n \exists y_1 \cdots \exists y_m \psi ,$$

où $\psi$ est une formule sans quantificateurs, dont les variables sont $x_1, \ldots, x_n$ et $y_1, \ldots, y_m$. Beaucoup de théories usuelles sont inductives, et toute théorie peut être "étendue" en une théorie inductive. On a les théorèmes suivants.

**Théorème 9.19.** La théorie $T$ est inductive si et seulement si, pour toute suite $\mathcal{M}_0, \mathcal{M}_1 \ldots$ de modèles de $T$ telle que $\mathcal{M}_i \subset \mathcal{M}_{i+1}$ pour tout $i$, la réunion $\bigcup_i \mathcal{M}_i$ est encore un modèle de $T$.

**Théorème 9.20.** Si $T$ est une théorie de langage $\mathcal{L}$, alors il existe une théorie $T'$ de langage $\mathcal{L}'$ telle que $T \subset T'$, $\mathcal{L} \subset \mathcal{L}'$ et que pour tout modèle $\mathcal{M}$ de $T$, il existe une expansion $\mathcal{M}'$ de $\mathcal{M}$ qui est un modèle de $T'$.

### 9.4.4. *Classification des théories*

Il existe des relations simples entre les concepts de complétude et de catégoricité introduits au paragraphe précédent. Ces relations se traduisent par quelques théorèmes.

**Théorème 9.21.** Une théorie catégorique est aussi $\alpha$-catégorique pour tout cardinal $\alpha$.

*Démonstration.* Il suffit d'observer que deux modèles isomorphes ont, par définition, le même cardinal.

*Remarque.* La réciproque de ce théorème est fausse; une théorie peut être $\alpha$-catégorique pour tout cardinal $\alpha$ et ne pas être catégorique. C'est le cas de la théorie de l'égalité. La signature de cette théorie élémentaire est vide, ce qui signifie que son langage comprend les formules ne comportant ni constantes fonctionnelles, ni constantes prédicatives (sauf l'égalité, qui appartient au lexique logique du calcul des prédicats avec identité). Des exemples de telles formules sont

(9.44) $\forall x\,(x = x)$,
$\forall x\,\exists y\,(x = y)$,
$\forall x\,\forall y\,(x = y)$.

Les deux premières formules sont valides; la troisième est consistante mais non valide. Par exemple, elle est vraie dans le modèle $\{a\}$ et fausse dans le modèle $\{a, b\}$. Il est clair que tout ensemble non vide est un modèle de cette théorie; deux modèles sont isomorphes si et seulement s'ils sont *équipotents*, c'est-à-dire qu'ils ont même cardinal.

La propriété de catégoricité est très attrayante, puisqu'elle permet de caractériser une structure par une théorie. En fait, cette propriété est trop forte : vu le théorème de Löwenheim et Skolem (cf. 8.2), le modèle d'une théorie catégorique est toujours fini. Ce fait restreint grandement l'utilité des théories catégoriques; en effet, une structure finie peut toujours être *décrite* exhaustivement, en principe du moins. L'axiomatisation n'ajoute rien à l'énumération pure et simple des vérités de la théorie.

Par contre, il existe des théories $\alpha$-catégoriques intéressantes. Par exemple, la théorie de l'ordre total, dense et non borné est $\omega$-catégorique, ce qui veut dire que tous ses modèles infinis dénombrables sont isomorphes[21]. A ce propos, on a le résultat suivant.

**Théorème 9.22.** (Vaught, 1954) Toute théorie qui est $\alpha$-catégorique pour un cardinal $\alpha$ infini et qui n'admet que des modèles infinis est complète.

*Démonstration.* On raisonne par contraposition. Soit $T$ une théorie consistante, non complète et n'admettant que des modèles infinis. Vu l'incomplétude, on peut écrire dans le langage de $T$ une formule $A$ telle que les théories $T \cup \{A\}$ et $T \cup \{\neg A\}$ soient toutes deux consistantes (les modèles sont nécessairement infinis). Par le théorème de Löwenheim et Skolem, chacune des théories admet un modèle de cardinal $\alpha$; comme $A$ est vrai dans l'un de ces modèles et pas dans l'autre, $T$ n'est pas $\alpha$-catégorique.

### 9.4.5. *La logique propositionnelle*

La métathéorie des modèles se simplifie fortement dans le cadre propositionnel. Dans ce contexte, étant donné un ensemble (dénombrable) $\mathcal{P}$ de propositions atomiques, une structure est une application de $\mathcal{P}$ dans $\{\mathbf{V}, \mathbf{F}\}$. Une théorie est simplement un ensemble de formules dont les propositions atomiques appartiennent à $\mathcal{P}$. Toute théorie complète est en même temps catégorique; en effet, si la théorie $T$ est complète, son unique modèle est déterminé par la règle suivante : $\mathcal{I}(p) = \mathbf{V}$ si et seulement si $p$ est un théorème de $T$. La théorie vide

[21] Le symbole $\omega$ désigne ici le cardinal de l'ensemble $\mathbf{N}$ des entiers naturels, et donc le cardinal de tout ensemble infini dénombrable.

est le calcul des propositions; ses théorèmes sont les tautologies. Cette théorie est incomplète. Plus précisément, aucun littéral n'est une tautologie.

Néanmoins, plusieurs résultats important de la métathéorie des modèles du premier ordre restent intéressants dans le cadre de la logique des propositions. Citons notamment le théorème d'interpolation[22] de Craig.

**Théorème 9.23.** Soient deux formules $\phi$ et $\psi$ (de la logique du premier ordre) telles que $\phi \models \psi$. Il existe une formule $\theta$ dont tous les symboles extralogiques sont communs à $\phi$ et $\psi$, et telle que $\phi \models \theta$ et $\theta \models \psi$.

Dans le cas du calcul des propositions, ce théorème reste valable et permet d'obtenir une forme simple d'une combinaison booléenne $\theta$ vérifiant la double contrainte $\phi \supset \theta \supset \psi$. La démonstration est assez difficile dans le contexte du premier ordre mais élémentaire dans le cas du calcul des propositions.

### 9.4.6. *Applications*

La théorie des modèles est très riche et permet d'obtenir des résultats intéressants dans des domaines variés. Néanmoins, par sa généralité même, elle met en jeu des concepts nombreux et souvent difficiles. Le but de ce paragraphe est de rendre plus intuitifs certains concepts et d'en donner des applications plus concrètes. Quoique la théorie des modèles concerne au premier chef le calcul des prédicats, elle s'applique souvent avec profit au calcul des propositions; dans ce contexte simplifié, certains résultats sont plus faciles à comprendre.

Le théorème de compacité est à la base de la théorie des modèles, aussi importe-t-il de bien en évaluer la portée. Le terme "compacité" provient de la topologie, et notre première application consistera à expliquer le lien entre le théorème de compacité et un certain espace topologique. (Aucune notion préalable de topologie n'est requise pour comprendre le paragraphe suivant.)

#### 9.4.6.1. *Théorème de compacité et topologie*

Soit $E$ un ensemble et $2^E$ l'ensemble des parties de $E$. Un sous-ensemble $\mathcal{T}$ de $2^E$ est une *topologie* sur $E$ s'il contient les unions finies et les intersections quelconques de ses éléments (en particulier, $\emptyset \in \mathcal{T}$ et $E \in \mathcal{T}$, car $\emptyset$ et $E$

[22] Hodges concrétise ce théorème en ces termes :

> Le théorème d'interpolation dit, *grosso modo*, que si un raisonnement valide a une prémisse unique qui mentionne la sauge mais pas les oignions et une conclusion qui mentionne les oignions mais pas la sauge, alors il peut être subdivisé de manière telle que le pas intermédiaire ne mentionne ni les oignions ni la sauge (Hodges, 1977, p. 142).

sont respectivement l'union et l'intersection de 0 élément de $\mathcal{T}$). Si $\mathcal{T}$ est une topologie sur $E$, le couple $(E, \mathcal{T})$ est un *espace topologique*; les éléments $F \in \mathcal{T}$ sont les (sous-ensembles) *fermés*[23] de $E$.

Un sous-ensemble $\Lambda$ de $\mathcal{T}$ a la *propriété d'intersection finie* si l'intersection d'un nombre fini d'éléments de $\Lambda$ est non vide. L'espace topologique $(E, \mathcal{T})$ est *compact* si tout sous-ensemble $\Lambda$ de $\mathcal{T}$ ayant la propriété d'intersection finie a une intersection $\bigcap \Lambda$ non vide.

*Remarque.* Ces définitions sont de prime abord un peu arides, mais ont des applications très concrètes. Par exemple, l'intersection d'une famille d'intervalles fermés emboîtés et non vides n'est jamais vide dans l'ensemble $\mathbf{R}$ des réels mais peut l'être dans l'ensemble $\mathbf{Q}$ des rationnels. Pour cette raison, un intervalle fermé borné de réels est dit compact, contrairement à un intervalle fermé borné de rationnels.

Supposons que $P$ soit un ensemble (infini) de propositions. L'ensemble $E =_{def} 2^P$ des parties de $P$ s'identifie à l'ensemble des interprétations (cf. section 3.4.3). On définit un sous-ensemble $\mathcal{T}$ de $E$ en spécifiant qu'un ensemble d'interprétations $\Lambda$ appartient à $\mathcal{T}$ si et seulement s'il existe un ensemble $\Phi$ de formules tel que $\Lambda$ soit l'ensemble de tous les modèles de $\Phi$. Ceci donne lieu à la proposition suivante.

**Lemme 9.24.** L'ensemble $\mathcal{T}$ est une topologie pour $E$.

*Démonstration.* Il est clair que $\mathcal{T}$ contient $E$ et $\emptyset$, qui sont respectivement l'ensemble des modèles des formules $(p \vee \neg p)$ et $(p \wedge \neg p)$. Soit maintenant un sous-ensemble $\Lambda \subset \mathcal{T}$. On vérifie facilement que $\bigcap \Lambda \in \mathcal{T}$; en effet, si, pour tout $\lambda \in \Lambda$, on note $\Phi_\lambda$ un ensemble de formules dont $\lambda$ constitue l'ensemble des modèles, alors $\bigcap \Lambda$ est l'ensemble des modèles de $\bigcup_{\lambda \in \Lambda} \Phi_\lambda$.
Il reste à vérifier que $\mathcal{T}$ contient les unions finies de ses éléments. Soient $\Phi_i, i = 1, \ldots, n$, des ensembles de formules dont $\lambda_i, i = 1, \ldots, n$ sont les ensembles de modèles. L'ensemble $\Phi = \{(\phi_1 \vee \cdots \vee \phi_n) : \phi_i \in \Phi_i \text{ pour tout } i = 1, \ldots, n\}$ admet le modèle $A$ si et seulement si $A \in (\lambda_1 \cup \cdots \cup \lambda_n)$. Pour $n = 2$, cette proposition découle du fait que tous les énoncés du développement suivant sont

[23] Notre définition n'est pas classique. La manière classique de procéder est d'introduire les notions de topologie au moyen des ensembles *ouverts*, qui sont les complémentaires des ensembles fermés. Les définitions moins classiques données ici sont équivalentes, et plus commodes dans le cas présent.

simultanément vrais ou faux :

$A \in (\lambda_1 \cup \lambda_2)$,

$A \in \lambda_1$ ou $A \in \lambda_2$,

pour tout $\phi_1 \in \Phi_1$, on a $A \models \phi_1$ ou pour tout $\phi_2 \in \Phi_2$, on a $A \models \phi_2$,

pour tout $\phi_1 \in \Phi_1$, pour tout $\phi_2 \in \Phi_2$, on a $A \models \phi_1$ ou on a $A \models \phi_2$,

pour tout $\phi_1 \in \Phi_1$, pour tout $\phi_2 \in \Phi_2$, on a $A \models (\phi_1 \vee \phi_2)$,

pour tout $\phi \in \Phi$, on a $A \models \phi$,

Le cas général se déduit facilement du cas $n = 2$.

Dans l'espace topologique $(E, \mathcal{T})$, on convient de désigner un ensemble fermé (un élément de $\mathcal{T}$) par $\lambda$ ou $\lambda_r$; un ensemble de formules dont ce fermé est l'ensemble des modèles sera noté $\Phi$ ou $\Phi_r$. Remarquons que $\Phi$ détermine $\lambda$ mais que la réciproque est fausse[24]. On a $\lambda = \{A : A \models \Phi\}$. Le lemme suivant établit un lien entre les notions de propriété d'intersection finie et de f-consistance.

**Lemme 9.25.** Dans l'espace topologique $(E, \mathcal{T})$, l'ensemble $\{\lambda_k : k \in K\}$ a la propriété d'intersection finie si et seulement si l'ensemble $\bigcup_{k \in K} \Phi_k$ est f-consistant.

*Démonstration.* Il suffit d'observer que, pour tout sous-ensemble fini $K_0 \subset K$, l'ensemble $\bigcap_{k \in K_0} \lambda_k$ est non vide si et seulement si l'ensemble $\bigcup_{k \in K_0} \Phi_k$ est consistant.

Les lemmes 9.24 et 9.25 permettent de déduire du théorème de compacité (9.9) le théorème suivant.

**Théorème 9.26.** L'espace topologique $(E, \mathcal{T})$ est compact.

*Démonstration.* Soit $\{\lambda_k : k \in K\}$ une famille de fermés ayant la propriété d'intersection finie. Vu le lemme 9.25, l'ensemble $\bigcup_{k \in K} \Phi_k$ est f-consistant. D'après le théorème de compacité, cet ensemble admet un modèle $A$, et on a nécessairement $A \models \Phi_k$, donc $A \in \lambda_k$, pour tout $k \in K$.

En fait, le théorème 9.26 est équivalent au théorème de compacité. Soit $\Psi = \{\psi_k : k \in K\}$ un ensemble f-consistant de formules; si on pose $\Phi_k = \{\psi_k\}$, on a $\Psi = \bigcup_{k \in K} \Phi_k$. Vu le lemme 9.25, l'ensemble $\{\lambda_k : k \in K\}$ a la propriété d'intersection finie. La compacité de l'espace topologique $(E, \mathcal{T})$ implique l'existence d'un modèle $A$ commun à tous les ensembles $\lambda_k$; $A$ est un modèle de $\Psi$ qui est donc consistant.

[24] Si on souhaite établir une correspondance biunivoque, il suffit de ne considérer que les ensembles de formules contenant toutes les conséquences logiques de leurs éléments.

*Remarque.* Les développements qui viennent d'être présentés pour le calcul des propositions s'adaptent facilement au calcul des prédicats.

### 9.4.6.2. *Théorème de compacité et théorie de l'ordre*

Le théorème de compacité permet de généraliser du cas fini au cas infini une propriété logique, la consistance. Dans la mesure où bien des propriétés mathématiques peuvent se formaliser dans le cadre de la logique du premier ordre, le théorème de compacité permet de généraliser au cas infini bien des propriétés dont le caractère logique n'apparaît pas immédiatement. La notion d'ensemble ordonné donne lieu, une fois de plus, à un exemple intéressant.

**Lemme 9.27.** Tout ensemble fini peut être muni d'une relation d'ordre total.

*Démonstration.* Chaque énumération d'un ensemble fini détermine un ordre total sur cet ensemble. Un ensemble à $n$ éléments admet $n!$ relations d'ordre total.

**Théorème 9.28.** Tout ensemble peut être muni d'une relation d'ordre total.

*Démonstration.* Soit $E$ un ensemble. Soit $\mathcal{L}$ le langage dont la signature comporte le prédicat binaire $\leq$, ainsi qu'une constante individuelle pour chaque élément $c \in E$. (Le même symbole est utilisé pour l'élément de $E$ et la constante associée.) On considère l'ensemble $T$ des énoncés $a \leq a$, $(a \leq b \wedge b \leq a) \supset a = b$ et $(a \leq b \wedge b \leq c) \supset a \leq c$, pour tout $a, b, c \in E$. Comme tout sous-ensemble fini de $E$ peut être totalement ordonné (lemme 9.27), l'ensemble $T$ est f-consistant, et donc aussi consistant.

*Remarques.* On peut imposer à l'ordre total de respecter un ordre partiel préexistant sur $E$; il suffit d'ajouter à l'ensemble $T$ les formules de type $a \leq b$ décrivant cet ordre partiel.
On verra au chapitre suivant (volume II) que l'axiome du choix (cf. 3.6.2) est un renforcement de l'énoncé 9.28.

### 9.4.6.3. *Le théorème des quatre couleurs*

Le théorème des quatre couleurs affirme qu'il est toujours possible de colorier une carte plane au moyen de quatre couleurs seulement, de telle sorte que deux régions ayant une frontière en commun (non limitée à un point) soient de couleurs différentes.

Ce résultat conjecturé depuis des siècles n'a été démontré que récemment, et sa preuve est longue et difficile. Sa validité dépend de la topologie du support de la carte; quatre couleurs suffisent sur un rectangle, un cercle, une sphère

ou un plan infini, mais ne suffisent plus sur des surfaces topologiquement plus complexes telles le tore, le ruban de Möbius ou la bouteille de Klein.

Notre but dans ce paragraphe sera de déduire le cas du plan infini de celui du rectangle, supposé déjà démontré. Le plan étant assimilable à un "rectangle infini", on pressent que le théorème de compacité pourrait s'appliquer.

La forme précise des différentes régions d'une carte n'a pas d'importance pour le problème qui nous occupe. Il suffit de savoir si deux régions distinctes ont une frontière en commun ou non. (Le cas d'une frontière réduite à un point est exclu : on se rend facilement compte qu'un nombre illimité de régions peuvent se toucher en un même point.) On représentera donc une carte par un graphe dont chaque sommet représente une région. Deux sommets sont reliés par une arête si et seulement si les régions correspondantes ont une frontière commune. A chaque sommet est associé la couleur de la région correspondante et deux sommets unis par une arête ont des couleurs distinctes. Une carte plane finie sera représentée par un graphe planaire[25] fini, puisqu'on suppose, bien naturellement, que chaque région a une aire supérieure à une constante $\varepsilon$ fixée.

Le problème est maintenant de prouver que tout graphe planaire $\Gamma$ peut être colorié au moyen de quatre couleurs, en sachant que cette propriété est vraie pour tous les graphes planaires finis, et en particulier pour tous les sous-graphes finis, nécessairement planaires, de $\Gamma$.

*Remarque.* Si toute carte rectangulaire est coloriable en quatre couleur, toute carte plane finie l'est également, quelle que soit sa forme. En effet, toute carte plane finie à $N$ régions peut être incluse dans une carte rectangulaire à $N+1$ régions, et un coloriage adéquat de cette dernière induit un coloriage adéquat de la carte initiale.

Le problème de l'existence d'un coloriage en quatre couleurs pour un graphe donné $\Gamma$ peut se ramener au problème de la consistance d'un ensemble $E_\Gamma$ de formules. Ces formules appartiennent à un langage $\mathcal{L}$ dont la signature comporte une constante individuelle pour chaque sommet du graphe et quatre prédicats unaires $C_1, C_2, C_3, C_4$ correspondant aux quatre couleurs. L'ensemble $E_\Gamma$ se compose, pour chaque sommet $s$, d'une formule $F(s)$ exprimant que $s$ a exactement une couleur et, pour chaque arête $\{s, s'\}$, d'une formule $G(s, s')$ affirmant que $s$ et $s'$ n'ont pas la même couleur. On a donc $E = \{F(s) : s \in S_\Gamma\} \cup \{G(s, s') : \{s, s'\} \in A_\Gamma\}$, où

$$F(s) =_{def} C_1(s) + C_2(s) + C_3(s) + C_4(s) = 1, \quad \text{pour tout sommet } s, \tag{9.45}$$

[25] C'est-à-dire un graphe représentable sur un plan, sans croisement d'arêtes.

$$(9.46)\quad G(s,s') =_{def} (C_1(s) \wedge C_1(s')) + (C_2(s) \wedge C_2(s')) + (C_3(s) \wedge C_3(s')) + (C_4(s) \wedge C_4(s')) = 0, \text{ pour toute arête } \{s,s'\},$$

où $S_\Gamma$ et $A_\Gamma$ sont respectivement l'ensemble des sommets et l'ensemble des arêtes de $\Gamma$. Il est clair que l'ensemble $E$ est consistant si et seulement si $\Gamma$ peut être colorié en quatre couleurs. D'autre part, le graphe $\Gamma$ est fini si et seulement si l'ensemble $E_\Gamma$ est fini.

*Remarque.* On utilise ici l'addition booléenne, qui se réduit à l'addition classique si on assimile **F** à 0 et **V** à 1; la formule $a_1 + a_2 + \cdots + a_n = k$ est vraie si et seulement si exactement $k$ des formules $a_1, \ldots, a_n$ sont vraies.

On veut prouver que $\Gamma$ est coloriable (en quatre couleurs) si ses sous-graphes finis sont coloriables. Comme la coloriabilité d'un sous-graphe fini de $\Gamma$ équivaut à la consistance d'un sous-ensemble fini de $E_\Gamma$ (obtenu en supprimant toutes les formules évoquant un sommet n'appartenant pas au sous-graphe), il est suffisant d'établir que $E_\Gamma$ est consistant si tous ses sous-ensembles finis sont consistants. Ce dernier point est une conséquence immédiate du théorème de compacité, ce qui établit le résultat.

### 9.4.6.4. *Synthèse de circuits logiques*

Les ordinateurs et beaucoup d'autres appareils électroniques comportent des circuits logiques, qui réalisent des fonctions booléennes, c'est-à-dire des fonctions de $\{\mathbf{V},\mathbf{F}\}^n$ dans $\{\mathbf{V},\mathbf{F}\}$. Concrètement, un tel circuit comportera $n$ entrées et une sortie. A chaque entrée est appliquée une tension haute (pour **V**) ou basse (pour **F**), ce qui provoque à la sortie l'apparition d'une tension haute ou basse.

Un problème classique en synthèse de circuits logiques consiste à construire un circuit réalisant une certaine fonction $f(p_1, \ldots, p_n)$ au moyen de circuits élémentaires réalisant des fonctions élémentaires comme la conjonction et la négation. On résout le problème en réduisant la formule représentant $f$ à une formule équivalente dont les seuls connecteurs sont, par exemple, la conjonction et la négation.

Bien souvent, le circuit à construire n'est pas spécifié complètement. Au cas où $f$ est spécifié par deux implications $(f_1 \supset f)$ et $(f \supset f_2)$, on peut utiliser le théorème de Craig pour obtenir un choix acceptable de $f$ comportant aussi peu d'arguments que possible. Cela conduira, vraisemblablement, à un circuit aussi simple que possible.

Cette technique est intéressante même si elle n'est pas optimale et, à titre d'application, nous allons montrer comment on peut la mettre en pratique. Cela donnera lieu à une démonstration constructive du théorème de Craig.

Soit $X(f_1, f_2)$ la formule

(9.47) $\quad (f_1 \supset f) \wedge (f \supset f_2)$.

Cette formule est vraie si et seulement si elle est vraie quand $p$ est vrai et quand $p$ est faux ($p$ désigne ici une proposition quelconque). Une formule $A$ est vraie quand $p$ est vrai (faux) si et seulement si la formule $A[p/\mathbf{V}]$ (la formule $A[p/\mathbf{F}]$) est vraie[26].

On voit donc que si $u$ est une solution de $X(f_1[p/\mathbf{V}], f_2[p/\mathbf{V}])$ et si $v$ est une solution de $X(f_1[p/\mathbf{F}], f_2[p/\mathbf{F}])$, alors

(9.48) $\quad f =_{def} ((u \wedge p) \vee (v \wedge \neg p))$

est une solution de $X(f_1, f_2)$ : il suffit d'observer que $f[p/\mathbf{V}] \equiv u$ et $f[p/\mathbf{F}] \equiv v$.

Ce résultat est intéressant dans le cas où $p$ intervient à la fois dans $f_1$ et $f_2$. Il permet de réduire le problème général de résolution de l'inéquation $X(f_1, f_2)$ au cas particulier où $f_1$ et $f_2$ n'ont aucune proposition en commun. Dans ce cas particulier, la condition $(f_1 \supset f_2)$ implique l'inconsistance de $f_1$ ou la validité de $f_2$ (le lecteur est invité à démontrer ce résultat). Une solution de l'inéquation $X(f_1, f_2)$ est $\mathbf{F}$ si $f_1$ est inconsistant et $\mathbf{V}$ si $f_2$ est valide.

A titre d'exemple, considérons le cas de l'inéquation

(9.49) $\quad X(p \wedge q\,,\ (p \vee q) \wedge (r \supset p))$.

La réduction par rapport à $p$ conduit au système d'inéquations

(9.50) $\quad X(q, \mathbf{V})\,,\ X(\mathbf{F}, q \wedge \neg r)$.

dont les solutions sont respectivement $\mathbf{V}$ et $\mathbf{F}$. Une solution du système initial est donc $(\mathbf{V} \wedge p) \vee (\mathbf{F} \wedge \neg p)$, qui se réduit à $p$.

## 9.5. Décidabilité

La propriété de complétude, quoique moins forte que celle de catégoricité, reste cependant fort restrictive. Un concept tout aussi intéressant, et un peu moins restrictif, est celui de décidabilité, que nous allons introduire.

Une théorie est *décidable* si ses théorèmes forment un sous-ensemble récursif[27] du langage, c'est-à-dire s'il existe un algorithme permettant de décider si une formule est ou n'est pas un théorème. On a le résultat suivant.

**Théorème 9.29.** Toute théorie complète est décidable.

*Démonstration.* Soit $T$ une théorie complète. Pour toute formule $A$ appartenant au langage de $T$, un et un seul des ensembles $T \cup \{A\}$ et $T \cup \{\neg A\}$ sera

[26] La formule $A[p/B]$ est obtenue en remplaçant dans la formule $A$ toutes les occurrences de la proposition $p$ par la formule $B$.

[27] Rappelons qu'un sous-ensemble $A$ d'un ensemble $E$ est *récursif* s'il existe un algorithme permettant de décider, pour tout élément $x \in E$, si $x \in A$ ou si $x \in E \setminus A$. La théorie de la décidabilité est aussi appelée théorie de la récursion; il s'agit en fait d'une métathéorie.

inconsistant. La méthode de résolution permet de détecter lequel, et donc de décider, en un temps fini, si $A$ appartient à la théorie.

*Remarque.* La réciproque est fausse; des contre-exemples seront donnés plus loin.

La notion de décidabilité ne s'applique pas uniquement à des théories mais aussi à des *problèmes formels*. Un problème formel se compose d'un ensemble infini dénombrable d'objets formels, appelés *instances*, et d'une question dichotomique s'appliquant aux instances. Un exemple bien connu est le *problème de l'arrêt*. Ses instances sont tous les programmes syntaxiquement corrects que l'on peut écrire dans un langage formel donné[28]. La question est : "ce programme se termine-t-il quelles que soient les valeurs initiales attribuées aux variables ?".

La métathéorie de la récursion a pour objet l'étude des problèmes formels. Un problème est *décidable* si l'on peut programmer une technique permettant de déterminer en un temps fini la réponse à la question, pour n'importe quelle instance.

Dans le cadre de la logique propositionnelle, la méthode des tables de vérité est programmable. Ceci montre que le problème de la validité des formules de la logique propositionnelle est décidable. L'ensemble des instances est l'ensemble des formules bien formées et la question est : "cette formule est-elle une tautologie ?".

Dans le cadre de la logique prédicative, la situation est plus complexe. A toute théorie correspond naturellement un problème formel dont les instances sont les formules du langage de la théorie, et dont la question est : "cette formule est-elle un théorème de la théorie ?". La théorie est décidable si et seulement si le problème formel qui lui est associé est décidable.

Beaucoup de théories importantes sont indécidables; telles sont notamment le calcul des prédicats lui-même, la théorie des nombres, la théorie des ensembles et la théorie des groupes. Néanmoins, certaines théories voisines, obtenues en modifiant quelque peu le langage et/ou l'ensemble des axiomes, sont décidables. C'est le cas du calcul des prédicats monadiques (seuls les prédicats à une place sont admis), de la théorie additive des nombres (la multiplication est exclue) et de la théorie des groupes abéliens.

Les théories du premier ordre — et le calcul des prédicats est une théorie du premier ordre — sont *partiellement décidables* dans le sens suivant. Il est possible (en utilisant la méthode de résolution par exemple) de vérifier mécaniquement qu'une formule est un théorème. Le point gênant est que la méthode de résolution, appliquée à une formule qui n'est pas un théorème, ne fournit pas de réponse.

Le problème de l'arrêt, lui aussi, n'est que partiellement décidable; il en est

[28] L'un de ces langages est celui des *fonctions récursives* de Gödel.

de même pour le célèbre "dixième problème" de Hilbert. Ce problème a pour instances les équations polynomiales à coefficients entiers (à plusieurs variables). La question est : "cette équation admet-elle une solution entière?". Les recherches nombreuses dont ce problème a fait l'objet ont trouvé leur couronnement en 1970, quand Matiasijevič a établi l'inexistence d'un algorithme adéquat.

Les problèmes décidables font l'objet de la (méta)théorie de la *complexité*. En effet, bien des problèmes admettent des techniques de décision simples mais peu efficaces. Par exemple, la méthode des tables de vérité est pratiquement impossible à utiliser pour les longues formules, mettant en jeu plusieurs centaines de variables. De même, quoiqu'il soit théoriquement très simple de déterminer si un nombre naturel est premier (par la méthode du crible d'Eratosthène, par exemple), le problème devient pratiquement insoluble pour les nombres de plusieurs centaines de chiffres décimaux. Des techniques plus efficaces que la méthode des tables de vérité et que le crible d'Eratosthène ont été mises au point mais, fondamentalement, une des questions ouvertes les plus importantes de l'informatique concerne l'existence de méthodes efficaces pour la résolution de ces problèmes et d'autres du même genre.

## 9.6. Structures algébriques usuelles

L'étude des structures algébriques usuelles permet d'illustrer les concepts théoriques introduits dans ce chapitre et de montrer leur intérêt. On voit également que la logique et l'algèbre sont étroitement liées : d'une part, les développements de la logique formelle reposent directement sur l'algèbre (élémentaire et supérieure) et, d'autre part, la logique permet d'obtenir des résultats algébriques intéressants. (Seuls des résultats élémentaires seront présentés ici, sans démonstration.)

La notion de groupe est centrale en algèbre et intervient dans de nombreux domaines des mathématiques. Deux groupes distincts peuvent présenter des propriétés fort différentes et il n'est pas surprenant que la théorie des groupes soit incomplète et même indécidable. Il n'est pas possible non plus d'axiomatiser la notion de groupe fini : toute formule du premier ordre qui est vraie dans tous les groupes finis sera également vraie dans certains groupes infinis.

Les mathématiciens algébristes ont naturellement cherché à mieux appréhender le concept de groupe et, en particulier, à classer les groupes en fonction de propriétés additionnelles qu'ils peuvent avoir. Comme on l'a vu à propos de la théorie de l'ordre, cette démarche est également indiquée du point de vue du logicien. Le point remarquable est que l'approche algébrique et l'approche logique convergent largement : les propriétés qui se sont avérées intéressantes du point de vue algébrique donnent également lieu à des résultats intéressants du point de vue logique. L'exemple le plus significatif est formalisé par le théorème suivant.

**Théorème 9.30.** La théorie des groupes abéliens est décidable.

*Remarque.* Dans la théorie des groupes abéliens, il est d'usage de noter additivement l'opération et aussi de désigner le neutre par 0. L'ensemble des postulats de cette théorie sera donc

$$\begin{aligned}(9.51)\quad & \forall x \forall y \forall z\,((x+y)+z = x+(y+z)),\\ & \forall x\,(x+0=x),\\ & \forall x \exists y\,(x+y=0),\\ & \forall x \forall y\,(x+y=y+x).\end{aligned}$$

La théorie des groupes abéliens n'est toutefois pas complète. Par exemple, la formule $\forall x\,(x+x=0)$ est vraie dans certains groupes abéliens et fausse dans d'autres. La méthode utilisée pour étudier la théorie de l'ordre linéaire dense s'applique également à la théorie des groupes abéliens (mais de manière beaucoup plus compliquée). En particulier, il est possible de déterminer l'ensemble des complétions de cette théorie. Pour ce faire, on peut utiliser comme discriminant la formule donnée plus haut ou, plus généralement, la définition suivante. Si $p$ est un nombre premier, on dit qu'un groupe abélien est *d'ordre* $p$ s'il vérifie la formule suivante[29]:

$$(9.52)\quad \forall x\,(px=0),$$

où $px$ abrège l'écriture $x+x+\cdots+x$ de la somme de $p$ termes égaux à $x$. Un groupe abélien est *d'ordre infini*, ou *sans torsion*, si la somme d'un nombre quelconque de termes égaux (différents de 0) ne donne jamais 0 (on obtient donc un ensemble de postulats pour la théorie des groupes sans torsion en ajoutant l'infinité d'axiomes du type $\forall x\,(x \neq 0 \supset px \neq 0)$, pour tout nombre premier $p$). On a alors le résultat suivant.

**Théorème 9.31.** Si $p$ est un nombre premier, la théorie des groupes abéliens d'ordre $p$ est $\alpha$-catégorique pour tout cardinal $\alpha$.

Ce théorème ne s'étend pas aux groupes sans torsion. On dit qu'un groupe sans torsion est *divisible* si, pour tout entier positif $n$, la formule $\forall x \exists y\,(ny = x)$ est vraie. Le groupe numérique $(\mathbf{Q},+)$ est divisible, alors que $(\mathbf{Z},+)$ ne l'est pas. Cette distinction donne lieu au théorème suivant.

**Théorème 9.32.** La théorie des groupes abéliens sans torsion et divisibles est $\alpha$-catégorique pour tout cardinal $\alpha$ infini non dénombrable.

*Remarques.* Il est clair qu'un groupe abélien sans torsion est toujours infini; si $a$ est un élément d'un groupe abélien sans torsion, tous les termes $a$, $a+a$, $a+a+a$, ... sont distincts, sauf si $a=0$. D'autre part, il existe des groupes abéliens dénombrables, sans torsion et divisibles, qui ne sont pas isomorphes.

[29] D'autres dénominations sont parfois utilisées pour cette propriété; d'autre part, pour certains auteurs, l'ordre d'un groupe est le cardinal de ce groupe.

On peut "perfectionner" la théorie des groupes en ajoutant des postulats, mais aussi en étendant le langage. On obtient un *anneau* à partir d'un groupe abélien $(G,+,0)$ en introduisant une constante fonctionnelle à deux places supplémentaire, la multiplication (notée par un point); on ajoute aussi les postulats

$$(9.53)\quad \begin{aligned} &\forall x \forall y \forall z\,((x.y).z = x.(y.z))\,,\\ &\forall x \forall y \forall z\,((x+y).z = (x.z)+(y.z))\,,\\ &\forall x \forall y \forall z\,(x.(y+z) = (x.y)+(x.z))\,. \end{aligned}$$

Si on ajoute en plus une constante individuelle, notée 1, et les postulats

$$(9.54)\quad \begin{aligned} &\forall x\,(x.1 = x = 1.x)\,,\\ &\forall x\,(x \neq 0 \supset \exists y\,(x.y = 1 = y.x))\,,\\ &\forall x \forall y\,(x.y = y.x)\,, \end{aligned}$$

on obtient successivement un *anneau unitaire*, un *corps* et un *champ*. (Si seul le deuxième postulat manque, on a un *anneau unitaire commutatif*; les champs sont les corps commutatifs.) La structure $(\mathbf{Z},+,0,.,1)$ est un anneau unitaire commutatif; les structures $(\mathbf{Q},+,0,.,1)$, $(\mathbf{R},+,0,.,1)$ et $(\mathbf{C},+,0,.,1)$ sont des champs.

L'intérêt de ces extensions en algèbre est bien connu; en logique, elles donnent également lieu à des résultats intéressants. On a en particulier le théorème suivant.

**Théorème 9.33.** La théorie des champs finis est décidable.

Rappelons que la notion de groupe fini n'est même pas axiomatisable. Alors qu'il existe une grande variété de groupes finis, les champs finis forment une famille beaucoup plus homogène. Un autre point par lequel la notion de corps se démarque de celle de groupe est le suivant.

**Théorème 9.34.** Tous les corps finis sont des champs (corps commutatifs).

Une notion importante en algèbre est celle de champ algébriquement clos. Un champ est *algébriquement clos* si toute équation polynomiale dont les coefficients appartiennent au champ admet une racine dans ce même champ. Le champ des nombres complexes est algébriquement clos, mais le champ des nombres réels ne l'est pas. Cette notion est également utile en logique, comme le montre le résultat suivant.

**Théorème 9.35.** La théorie des champs algébriquement clos est décidable.

En algèbre comme en logique, il est intéressant de raffiner encore la notion de champ, au moyen de la définition suivante. Un champ est *de caractéristique* $p$, pour un nombre premier $p$, s'il vérifie le postulat $p1 = 0$ ; il est *de caractéristique* 0 s'il vérifie les postulats $p1 \neq 0$, pour tous les nombres premiers $p$. On a alors le résultat suivant.

**Théorème 9.36.** La théorie des champs algébriquement clos de caractéristique

$p$ fixée est $\alpha$-catégorique pour tout cardinal $\alpha$ non dénombrable.

*Remarque.* Ce théorème montre que toute tentative de généraliser la notion de nombre complexe se solde inévitablement par la perte de certaines propriétés. Par exemple, l'ensemble des quaternions (qui est l'extension la plus féconde du champ des nombres complexes) est un corps non commutatif.

**Corollaire 9.37.** La théorie des champs algébriquement clos de caractéristique $p$ fixée est complète.

# Bibliographie


[1] Abramski et al., *Handbook of Logic in Computer Science* (à paraître).

[2] Ackermann W., voir Hilbert D.

[3] Anderson J. M., Johnstone H. W., *Natural Deduction*, Wadsworth, Belmont, 1962.

[4] Andrews P. B., *An Introduction to Mathematical Logic and Type Theory: to Truth through Proofs*, Academic Press, New York, 1986.

[5] Aristote, *Organon : I Catégories, II De l'interprétation*, trad. et notes de J. Tricot, Vrin, Paris, 1959.

[6] Aristote, *Organon : III Les premiers analytiques*, trad et notes de J. Tricot, Vrin, Paris, 1947.

[7] Aristote, *La métaphysique*, trad et notes de J. Tricot, Vrin, Paris, 1948.

[8] Arnauld A., Nicole P., *La logique ou l'art de penser*, ed. de P. Clair et F. Girbal, Paris 1965, 1981, 1$^{e}$ ed. 1662.

[9] Austin J. L., *Philosophical Papers*, édités par J. O. Urmson et G. J. Warnock, 1961, ed. augmentée, Oxford Univ. Press, 1970.

[10] Avron A., "Gentzenizing Schroeder-Heister's Natural Extension of Natural Deduction", LFCS Report Series, Edinburgh, 1987, p. 1–10.

[11] Ayres F., *Theory and Problems of Modern Algebra 1965*, Mc Graw Hill, Schaum, 1965, trad. sous le titre, *Algèbre Moderne*, Trad. de M. Lobenberg, Ediscienco, Paris, 1979.

[12] Barwise J. (ed.), *Handbook of Mathematical Logic*, North-Holland, Amsterdam, 1977.

[13] Beckman F., *Mathematical Foundations of Programming*, Addison-Wesley, Menlo Park, 1980.

[14] Bell J., Machover M., *A Course in Mathematical Logic*, North-Holland, Amsterdam, 1977.

[15] Belnap N., "Tonk, Plonk and Plink", *Analysis*, 1962, repris dans *Philosophical Logic*, edited by P. F. Strawson, Oxford Univ. Press, 1967, p. 132–137.

[16] Belnap N., "A useful four-valued Logic", in *Modern Uses of Multiple-Valued Logic*, édité par J.M. Dunn and G. Epstein, Reidel, Dordrecht, 1977, p. 8–37.

[17] Belnap N., "How a Computer should think", in Ed. G. Ryle, *Contemporary Aspects of Philosophy*, Oxford, 1976, p. 30–69.

[18] Beth E., *Aspects of Modern Logic*, Dordrecht, Reidel, 1970.

[19] Ben-Chorin (Mordechay) S., "Sentence Meanings as a Formal Basis for Intensional Logic", UCLA Working Papers in Semantics, Los Angeles, 1983, p. 86–117.

[20] van Benthem J., "The Variety of Consequence According to Bolzano", *Studia Logica*, XLIV, 1985, p. 321–351.

[21] Bergman E., Noll H., *Mathematische Logik mit Informatik-Anwendungen*, Springer-Verlag, Berlin, 1977.

[22] Besnard Ph., *An Introduction to Default Logic*, Spinger Verlag, 1989.

[23] Bibel W., *Automated Theorem Proving*, Vieweg, Braunschweig, 1982.

[24] Bibel W., "Perspective of Automated Deduction", draft, *10th Conference on Automated Deduction*, Kaiserlautern, 1990.

[25] Blanché R., *La logique et son histoire. D'Aristote à Russell*, Armand Colin, Paris, 1970.

[26] Bledsoe W. W., Loveland D. W. (ed.), *Automated Theorem Proving: After 25 Years*, American Mathematical Society, Providence, 1984.

[27] Bolzano B., *Wissenschaftslehre*, édité et trad. sous le titre *Theory of Science*, par R. George, Basil Blackwell, Oxford, 1972.

[28] Bridge J., *Beginning Model Theory*, Clarendon Press, Oxford, 1977.

[29] Carnap R., *The Logical Syntax of Language*, trad. de l'allemand par A. Smeaton, Routledge and Kegan Paul, Londres, 1937, 6$^{e}$ ed. 1964.

[30] Chang Ch. L., Lee R. Ch. L., *Symbolic Logic and Mechanical Theorem Proving*, Academic Press, New York, 1973.

[31] Chang C., Keisler H., *Model Theory*, North-Holland, Amsterdam, 1973.

[32] Chomsky N., *Aspects de la théorie syntaxique* (1965), trad. de l'anglais par Milner, Le Seuil, Paris, 1971.

[33] Church A., *Introduction to Mathematical Logic*, Princeton University Press, Princeton, 1956.

[34] Cocchiarella N., *Logical Investigations of Predication Theory and The Problem of Universals*, Bibliopolis, Naples, 1986.

[35] Cocchiarella N., "Predication Versus Membership in the Distinction Between Logic as Language and Logic as Calculus", *Synthese*, 77, 1988, p. 37–72.

[36] Copi I., *Symbolic Logic*, Colliers Mc Millan, Londres, 3e ed. 1967, 5e ed. 1986.

[37] Corcoran J., "Categoricity", *History and Philosophy of Logic*, 26, 1980, p. 187–207.

[38] Czelakowski J., Malinowski G., "Key Notions of Tarski's Methodology of Deductive Systems", *Studia Logica*, XLIV, 1985, p. 321–351.

[39] Davidson D., "True to the Facts", *The Journal of Philosophy*, 1967, repris dans *Inquiry into Truth and Interpretation*, Clarendon Press, Oxford, 1984.

[40] Delahaye J. P., *Outils logiques pour l'intelligence artificielle*, Editions Eyrolles, Paris, 1986.

[41] Delahaye J. P., "Nouvelles solutions pour les paradoxes sémantiques", Université des Sciences et Techniques de Lille, Laboratoire d'Informatique Fondamentale de Lille.

[42] Descartes R., *Discours de la méthode*, 1e ed. 1637, avec introduction et notes de E. Gilson, Vrin, Paris, 1954.

[43] Desclés J. P., "Réseaux sémantiques : la nature logique et linguistique des relateurs", *Langages*, 87, 1987, p. 55–78.

[44] Dubois D., Prade H., *Théorie des Possibilités. Applications à la représentation des connaissances en informatique*, Masson, Paris, 1988.

[45] Dubucs J. P., "Die sogennante Analyzität der Mathematik: Für eine Radikalisierung der Theorie Hintikkas", *Grazer Philosophische Studien*, 1988, 32, p. 83–112.

[46] Dummett M., *Frege, Philosophy of Language*, Duckworth, Londres, 1973, 2e ed. 1981.

[47] Eaton R., *General Logic. An Introductory Survey*, Charles Soabners's Sons, New York, 1e ed. 1931, 1959.

[48] Enderton H., *A Mathematical Introduction to Logic*, Academic Press, New York, 1972.

[49] Engel P., "La Norme du Vrai : Philosophie de la logique", Gallimard, Paris, 1989.

[50] Etchemendy J., "Tarski on Truth and Logical Consequence", *The Journal of Symbolic Logic*, 53, 1988, p. 51–79.

[51] Etchemendy J., "The Doctrine of Logic as Form", *Linguistics and Philosophy*, 1983, p. 319–334.

[52] Feferman S., *The Numbers Systems*, Addison-Wesley, Reading Mass., 1964.

[53] Fine K., *Reasoning with Arbitrary Objects*, Aristotelian Society Series, vol. 3, Blackwell, Oxford, 1985.

[54] Frege G., *Begriffsschrift (1879)*, traduit et annoté par J. van Heijenoort dans *From Frege to Gödel*, Harvard University Press, Harvard, 1967, p. 1–82.

[55] Frege G., *Ecrits logiques et philosophiques*, traduction et introduction de Cl. Imbert, le Seuil, Paris, 1971.

[56] Frege G., *Les fondements de l'arithmétique*, traduction et introduction de Cl. Imbert, le Seuil, Paris, 1969.

[57] Frost R. A., *Introduction to Knowledge Base Systems*, Collins, London, 1986, réimpr. 1987.

[58] Gallier J., *Logic for Computer Science, Foundations of Automatic Theorem Proving*, Harper and Row, New York, 1986.

[59] Galmiche M., *Introduction à la sémantique de R. Montague*, P. U. F., Paris, 1990.

[60] Gauthier Y., *Méthodes et concepts de la logique formelle*, Presses de l'Université de Montréal, 1978.

[61] Gentzen G., *The Collected Papers of Gerhard Gentzen*, édité par M. E. Szabo, North-Holland, Amsterdam, 1969.

[62] Gochet P., *Esquisse d'une théorie nominaliste de la proposition*, A. Colin, Paris, 1972.

[63] Gochet P., *Ascent to Truth*, Philosophia Verlag, Munich, 1986.

[64] Goldfarb W., "Logic in the Twenties : the Nature of the Quantifier", *The Journal of Symbolic Logic*, 44, 1979, p. 351–368.

[65] Granger G. G., *La théorie aristotélicienne de la science*, Aubier, Paris, 1976.

[66] Grize J. Bl., "Logique", 135–289, *Logique et connaissance scientifique*, *Encyclopédie de la Pleiade*, sous la direction de J. Piaget, Gallimard, Paris, 1967.

[67] Grosjean P. V., *Initiation aux mathématiques nouvelles à partir de problèmes techniques*, Dunod, Paris, 1970.

[68] Guenthner F., Sabatier P., "Sémantique formelle et représentation des connaissances", *Langages*, 87, 1987, p. 103–122.

[69] Haack S., *Philosophy of Logics*, Cambridge University Press, Cambridge, 1978.

[70] Hamilton A. G., *Logic for Mathematicians*, Cambridge U. P. 1e ed., 1978.

[71] van Heijenoort J., *From Frege to Gödel, A Source Book in Mathematical Logic*, 1879–1931, Harvard University Press, Harvard, 1967.

[72] van Heijenoort J., "Logic as Calculus and Logic as Language", *Synthese*, 17, 1967, p. 324–330.

[73] Hilbert D., Ackermann W., *Principles of Mathematical Logic*, trad. de l'allemand (2e ed. 1938), Chelsea Publ. Co., 1950.

[74] Hintikka J., *Logic, Language Games and Information*, Clarendon Press, Oxford, 1973.

[75] Hintikka J., "On the Development of the Model-Theoretic Viewpoint in Logical Therory", *Synthese*, 77, 1988, p. 1–36.

[76] Hirsch G., "Mathématisation et Réalité", *Dialectica*, 29, 1975, p. 5–24.

[77] Hodges W., *Logic*, Penguin Books, Harmondsworth, 1977.

[78] Hodges W., "Truth in a Structure", in *Proceedings of the Aristotelian Society*, 1985–1986, p. 135–151.

[79] Hubien H., "A New Basis for Classical Propositional Calculus", *Logique et analyse*, 79, 1977, p. 225–227.

[80] Hughes G. E., Londey D. G., *The Elements of Formal Logic*, Methuen, London, 1965.

[81] Imbert Cl., "Le projet idéographique", *Revue Internationale de Philosophie (Frege 1879–1979)*, 130, 1979, p. 621–665.

[82] Ishiguro H., *Leibniz's Philosophy of Logic and Language*, Londres, Duckworth, 1972.

[83] Jacques F., "L'idéographie frégéenne : un langage libéré des contraintes de l'interlocution", *Revue Internationale de Philosophie Frege (1879–1979)*, 130, 1979, p. 694–715.

[84] Jayez J., *L'inférence en langue naturelle*, Hermès, Paris, 1988.

[85] Johnstone H. W., voir Anderson J. M.

[86] Kant E., *Critique de la raison pure*,1e ed. 1781, trad. de l'allemand par A. Tremesaygues et B. Pacaud, P. U. F., Paris, 1950.

[87] Keisler H., voir Chang C.

[88] Kilmister C. W., *Language, Logic and Mathematics*, The English Universities Press Limited, Londres, 1967.

[89] Kirwan C., *Logic and Argument*, Londres, 1978.

[90] Kleene S. C., *Introduction to Metamathematics*, Elsevier North-Holland, New York, 1952.

[91] Kleene S. C., *Mathematical Logic*, 1967, traduit par A. Largeault sous le titre *Logique mathématique*, A. Colin, Paris, 1972.

[92] Kneale W. et M., *The Development of Logic*, Clarendon Press, Oxford, 1962.

[93] Kremer M., "Logic and Meaning : The Philosophical Significance of the Sequent Calculus", *Mind*, 97, 1988, p. 50–72.

[94] Ladrière J., "Le théorème fondamental de Gentzen", *Revue philosophique de Louvain*, 1951, p. 357–384.

[95] Ladrière J., "Le formalisme et le sens", *Les langages, le sens et l'histoire*, Colloque international à l'Université de Lille III, janvier–mai 1972, Publications de l'Université de Lille III, Lille, 1$^{e}$ ed. 1975, 2$^{e}$ ed. 1977, p. 241–277.

[96] Lalande A., *Vocabulaire technique et critique de la philosophie*, Presses Universitaires de France, Paris, 1951.

[97] Langford C. H., voir Lewis Cl. I.

[98] Largeault J., *Quine, Questions de mots, Questions de faits*, Privat, Toulouse, 1980.

[99] Laurière J. L., *Intelligence artificielle*, résolution de problèmes par l'Homme et la machine, Eyrolles, Paris, 1987, 2$^{e}$ ed. revue et corrigée.

[100] Leblanc H., Wisdom W., *Deductive Logic*, Allyn and Bacon, Boston, 1972.

[101] Lee R. Ch. L., voir Chang.

[102] Lewis H.R., Papadimitriou Ch. H., *Elements of the Theory of Computation*, Prentice-Hall, Englewood Cliffs, 1981.

[103] Londey D. G., voir Hughes G. E.

[104] Loveland D., *Automated Theorem Proving : a Logical Basis*, North-Holland, Amsterdam, 1978.

[105] Machover M., voir Bell J.

[106] Manna Z., *Mathematical Theory of Computation*, McGraw-Hill, New York, 1974.

[107] Martin-Löf P., "Truth of a proposition, evidence of a judgement, validity of a proof", *Synthese*, 73, 1987, p. 407–420.

[108] Massey G. J., *Understanding Symbolic Logic*, Harper and Row, New York, 1970.

[109] Martinez Freire P., *Introducción a la logica matemática*, Editorial Libreria Agora, Malaga, 1985.

[110] Mates B., *Elementary Logic*, Oxford U. P., 1965.

[111] Martin R., *Logique contemporaine et formalisation*, Presses Universitaires de France, Paris, 1964.

[112] Mendelson E., *Introduction to Mathematical Logic*, Van Nostrand, Princeton, 1963.

[113] Mill J. S., *A system of Logic*, 1943, traduit sous le titre de , *système de logique*, par Louis Peisse, Librairie philosophique de Ladrange, Paris, 1866, réédité par Pierre Mardaga, Liège, 1988.

[114] Mints G., "Gentzen-Type Systems and Resolution Rules", Part I, Propositional Logic *Colog*, Lecture Notes in Computer Science, vol. 417, 1990.

[115] Moens J. L., "Déduction naturelle", *Intuitionnisme et théorie de la démonstration, Cahiers du Centre de Logique*, I, Louvain, Cabay, 1980, p. 45–67.

[116] Moisil Gr. C., *Essais sur les logiques non chrysipiennes*, Editions de l'Académide de la République Socialiste de Roumanie, Bucarest, 1972.

[117] Montague R., *Formal Philosophy*, Selected Papers of R. Montague, édité avec une introduction par R. H. Thomason, Yale Univ. Press, Yale, 1974.

[118] Moore G. H., "The Emergence of First-Order Logic", *Minnesota Studies in the Philosophy of Science, Philosophy and History of Modern Mathematics* (W. Aspray et Ph. Kitcher ed.), 1988, vol. 11.

[119] Murawski R., "The Development of Symbolism in Logic and its philosophical Background", *Proceedings of the Aristotelian Society 1988/89*, LXXXIX, p. 65–78.

[120] Nef F., *Logique et Langage*, Essais de sémantique intensionnelle, Hermès, Paris, 1988.

[121] Noll H., voir Bergmann.

[122] Novikov P. S., *Introduction à la logique mathématique*, trad. du russe par Ch. Sarthou, Dunod, Paris, 1964.

[123] Papadimitriou Ch. H., voir Lewis H. R.

[124] Partee B., "Montague Grammar and Transformational Grammar", *Linguistic Inquiry*, 1975, p. 203–300.

[125] Partee B., *Fundamentals of Mathematics for Linguistics*, Reidel, Dordrecht, 1978.

[126] Pascal Bl., *Oeuvres complètes*, "De l'esprit géométrique", "De l'art de persuader" (rédigés en 1656–1657), Le Seuil, Paris, 1963.

[127] Piaget J. et al., *Logique et connaissance scientifique*, *Encyclopédie de la Pleiade*, Gallimard, Paris, 1967.

[128] Pire F., *Questions de psychologie*, De Boeck, Bruxelles, 1989.

[129] Platon, *Le sophiste*, in *Oeuvres Complètes* Tome VIII — 2$^{e}$ partie, texte établi et raduit par A. Diès, Les Belles Lettres, Paris, 1963.

[130] Poincaré H., *La science et l'hypothèse*, Paris 1902, réédition Flammarion, Paris, 1968.

[131] Popper K., "Philosophical Comments on Tarski's Theory of Truth", *Objective Knowledge*, Clarendon Press, Oxford, 1972.

[132] Prade H., voir Dubois D.

[133] Prawitz D., *Natural Deduction, A Proof-Theoretical Study*, Almqvist and Wiskell, Stockolm, 1965.

[134] Prawitz D., "Ideas and Results in Proof Theory", In J. E. Fenstad, Ed., *Proceedings of the Second Scandinavian Logic Symposium*, Amsterdam, 1971, p. 235–307.

[135] Prawitz D., "Toward a foundation of a general proof theory", In P. Suppes et al., Eds., *Logic, Methodology, and Philosophy of Science IV*, p. 225–250, Amsterdam, 1973.

[136] Prawitz D., "On the idea of a general proof theory", *Synthese*, 27, 1974, p. 63–77.

[137] Prawitz D., "Proofs and the Meaning of Completeness", in J. Hintikka, I. Niiniluoto and E. Saarinen (eds.), *Essay on Mathematical and Philosophical Logic*, 25–40, Reidel, Dordrecht, 1978.

[138] Prawitz D., "Philosophical aspects of proof theory", in *Contemporary philosophy, A new Survey* édité par G. Fløistad, Nijhoff, La Haye, 1981, p. 235–277.

[139] Prawitz D., "Remarks on Some Approaches to the Concept of Logical Consequence", *Synthese*, 62, 1985, p. 153–171.

[140] Putnam H., "A Comparison of Something with Something Else", *New Literary History*, 17, 1985, p. 61–79.

[141] de Queiroz J. G. B., "A Proof-Theoretic Account of Programming and The Role of Reduction Rules", *Dialectica*, 42, 1988, p. 265–282.

[142] Quine W.V.O., *Mathematical Logic*, Harvard Univ. Press, Harvard, 1940, édition revue, 9$^{e}$ impr, 1981.

[143] Quine W.V.O., *From a logical Point of View*, Logico-Philosophical Essays, 1953, 2$^{e}$ édition, Harper and Row, New York, 1961.

[144] Quine W.V.O., *Set Theory and its Logic*, The Belknap Press, Harvard, 1963.

[145] Quine W.V.O., *Méthodes de logique* (1$^{e}$ ed. 1958), trad. de l'anglais par M. Clavelin, P. U. F., Paris, 1973.

[146] Quine W.V.O., *Philosophy of Logic*, 1970, traduit par J. Largeault sous le titre de *Philosophie de la logique*, Aubier, Paris, 1975.

[147] Quine W.V.O., *Word and Object*, 1960, traduit par J. Dopp et P. Gochet sous le titre *Le mot et la chose*, Paris, Flammarion, 1978.

[148] Quine W.V.O., *The Roots of Reference*, Open Court, La Salle, 1973.

[149] Michael O. Rabin, "Decidable Theories", 515–629, in Jon Barwise, *Handbook of Mathematical Logic*, North-Holland, Amsterdam, 1977.

[150] Rasiowa H., *Introduction to Modern Mathematics*, North-Holland, Amsterdam, 1973.

[151] Reenfrank M. et al., *Non Monotonic Reasoning*, Springer Verlag, Berlin, 1986.

[152] Rivenc F., *Introduction à la logique*, Payot, Paris, 1989.

[153] Robinson J. A., A machine Oriented Logic Based on the Resolution Principle, 1965, repris dans *Automation and Reasoning*, Classical Papers on Computational Logic 1957–1966 édité par J. Siekmann and G. Wrightson, Springer-Verlag, Berlin, p. 397–415.

[154] Robinson J. A., *Logic : Form and Function. The Mechanization of Deductive Reasoning*, Edinburg University Press, Edimbourg, 1979.

[155] Rogers R., *Mathematical Logic and formalized theories*, North-Holland, Amsterdam, 1971.

[156] Runes D., *Dictionary of Philosophy*, Philosophical Library, New York, 15$^{e}$ ed., 1960.

[157] Russell, voir Whitehead.

[158] Russell B., *The Principles of Mathematics*, George Allen and Unwin, Londres, 1903, 2$^{e}$ ed. 1937, 8$^{e}$ impression 1964.

[159] Russell B., "On Denoting", 1905, in *Logic and Knowledge*, Essays 1901–1950 (M. Marsh, ed), Londres, George Allen & Unwin, 1956, trad. fr. dans *Philosophie*, 8, 1985, p. 11–28.

[160] Russell B., *Introduction to Mathematical Philosophy*, George Allen and Unwin, Londres, 1919.

[161] Russell B., *An Inquiry into Meaning and Truth*, George Allen and Unwin, Londres, 1940, 6$^{e}$ impr. 1966, traduit par Ph. Devaux sous le titre *Signification et vérité*, Flammarion, Paris, 1969.

[162] Ruyer B., *Logique*, P. U. F., Paris, 1990.

[163] Ryle G., "The Theory of Meaning", *British Philosophy in the Mid-Century* (C. A. Mace ed.), Londres, 1957, p. 239–264.

[164] Sabatier P., voir Guenthner F.

[165] Schroeder-Heister P., "A Natural Extension of Natural Deduction", *The Journal of Symbolic Logic*, 49, 1984, p. 1284–1300.

[166] Swichtenberg H., "Normalization", *Proceedings of the International Conference in Logic, Algebra and Computation*, ed. F. Bauer, Springer Verlag, 1990, à paraître.

[167] Swichtenberg H., "Proof Theory: Some Applications of Cut-Elimination", in *Handbook of Mathematical Logic* (R. Barwise ed.), North-

Holland, Amsterdam, 1977, p. 867–895.

[168] Scott D., "On engendering an illusion of Understanding", *The Journal of Philosophy*, 68, 1971, p. 787–807.

[169] Scott D., "Background to Formalization", *Truth, Syntax and Modality*, édité par H. Leblanc, North-Holland, Amsterdam, 1973, p. 244–273.

[170] Scott D., "Completeness and axiomatizability in many-valued logic", dans *Proceedings of the Tarski Symposium*, American Mathematical Society, Providence, 1974, p. 411–435.

[171] Scott D., "Rules and Derived Rules", *Logical Theory and Semantic Analysis* (S. Stenlund ed.), Reidel, Dordrecht, 1974, p. 147–161.

[172] Scott D. et al., *Notes on the Formalization of Logic*, Parts I & II, Parts III & IV, Oxford 1981, 2 vol.

[173] Shapiro S., "Second-Order Languages and Mathematical Practice", *The Journal of Symbolic Logic*, 50, 1985, p. 714–742.

[174] Sergant M., *Logisch redeneren, een inleiding in de logica*, Acco, Leuven, 1986.

[175] Shoenfield J., *Mathematical Logic*, Addison-Wesley, Menlo Park, 1967.

[176] Sluga H., *Gottlob Frege*, Routledge & Kegan Paul, Londres, 1980.

[177] Smart J. J. C., "How to Turn the *Tractatus* Wittgenstein into (Almost) Donald Davidson" in E. Lepore, *Truth and Interpretation*, Oxford, Blackwell, 1986, p. 92–100.

[178] Smets Ph. et al., *Non Standard Logic for Automated Reasoning*, Academic Press, Londres, 1988.

[179] Smiley, "The Schematic Fallacy", *Proceedings of the Aristotelian Society*, 1982–1983, p. 1–17.

[180] Smullyan R. M., *First Order Logic*, Springer-Verlag, New York, 1968.

[181] Sombé L., *Raisonnements sur des informations incomplètes en intelligence artificielle*, Teknea, Toulouse, 1989.

[182] Spinoza, "De la réforme de l'entendement" (1663), *Oeuvres Complètes*, traduit, présenté et annoté par Roland Caillois, Madeleine Francès et Robert Misrahi, Bibliothèque de la Pleiade, Gallimard, Paris, 1954.

[183] Stoll R., *Set Theory and Logic*, Freeman and Co., San Francisco, 1961.

[184] Strawson Sir P. F., *Introduction to Logical Theory*, Methuen, Londres, 1952.

[185] Sundholm G., "Systems of Deductions", in *Handbook of Philosophical Logic* Volume I, Elements of Classical Logic, édités par D. Gabbay and F. Guenthner, Reidel, Dordrecht, 1983, p. 133–188.

[186] Suppes P., "A Set of Independent Axioms for Extensive Quantities", 1951, repris dans *Studies in the Methodology and Foundations of Science*,

Reidel, Dordrecht, 1969, p. 36–45.

[187] Suppes P., *Introduction to Logic*, Van Nostrand, New York, 1957.

[188] Tarski A., *Introduction to Logic and the Methodology of Deductive Sciences*, Oxford University Press, Oxford, New York, 1e ed 1941, revised ed. 1946.

[189] Tarski A., Motowski A. et Robinson R. M., *Undecidable Theories*, North-Holland, Amsterdam, 1953.

[190] Tarski A., "Truth and Proofs", *L'Age de la Science*, 2, 1969, p. 279–301.

[191] Tarski A., "Le concept de vérité dans les langages formalisés", 157–269, *Logique, sémantique, métamathématique*, 1923–1944 tome 1, Armand Colin, Paris, 1972.

[192] Tarski A., "Sur le concept de conséquence logique", 141–152, *Logique, sémantique, métamathématique*, 1923–1944 tome 2, Armand Colin, Paris, 1972.

[193] Tennant N., *Natural Deduction*, Edinburgh University Press, Edimbourg, 1978.

[194] Tennant N., *Anti-Realism and Logic*, Truth as Eternal, Clarendon Press, Oxford, 1987.

[195] Thayse A. et al., *Approche logique de l'intelligence artificielle*, 1. De la logique classique à la programmation logique, 2. De la logique modale à la logique des bases de données, Dunod, Paris, Tome I 1988, tome II 1989.

[196] Vernant D., *Introduction à la philosophie de la logique*, Mardaga, Bruxelles, 1986.

[197] Vuillemin J., *Nécessité ou contingence*, Les Editions de Minuit, Paris, 1984.

[198] Wagner S., "The rationalist conception of logic", *Notre Dame Journal of Formal Logic*, 28, 1987, p. 3–35.

[199] Wallen L., *Automated Proof Search in Non-Classical Logics : Efficient Matrix Proof Methods for Modal and Intuitionistic Logics*, Edinburgh, Ph. D., Edimbourg, 1987, M.I.T. 1989.

[200] Watanabe S., "Une explication mathématique du classement d'objets", in *Information and Prediction in Science*, édité par S. Dockx et P. Bernays, Academic Press, New York and London, 1965, p. 39–76.

[201] Weingartner P., "A simple Proposal for introducing Intensions into Science", *La sémantique dans les Sciences*, Archives de l'Institut International des Sciences Théoriques, Office international de Librairie, Bruxelles, 1978, p. 193–209.

[202] Whitehead A., Russell B., *Principia Mathematica*, Cambridge Univ.

Press, Cambridge, 1962, (1e ed. des Principia Mathematica : 1913, 2e ed. : 1927).

[203] Winograd T., *Understanding Natural Language*, Academic Press, New York et Londres, 1972.

[204] Wisdom W., voir Leblanc H.

[205] Wittgenstein L., *Tractatus logico-philosophicus* with a new translation by D. F. Pears & B. F. Mc Guinness, Routledge & Keagan Paul, Londres, 1961 [1e ed. all. 1921].

[206] Zermelo E., "Sur les ensembles finis et le principe de l'induction complète", 1907, repris dans *Poincaré, Russell, Zermelo et Peano*, Textes réunis par G. Heinzmann, Blanchard, Paris, 1986.

[207] Zuber R., *Structure présuppositionnelle du langage*, Dunod, Paris, 1972.

[208] Zuber R., *Implications sémantiques dans les langues naturelles*, Editions du CNRS, Paris, 1990.

# Index


Absorbant, 46
Abstraction,
    principe d'—, 337
Actives *versus* passives,
    formules —, 250
    phrases —, 64
    propositions —, 241
Adéquation,
    conditions d'— matérielle et formelle d'une définition de la vérité, 334
Algèbre,
    — de Boole, 55
    — des classes, 55
    — ordinaire, 44, 90
Algorithme, 81
    — de construction de tableau, 245
    — de décision, 87
    — d'analyse de vérité (Quine), 100
    — de résolution, 285
    — d'unification (Robinson), 259, 272, 283
    — impraticable, 91
    — pour l'homme *versus* pour l'ordinateur, 193
Ambiguïté,
    — de $x$, 206, 224
    — inhérente à l'algèbre booléenne interprétée, 59
    — sémantique, 54
    — syntaxique, 54
Ancêtre, 360
    — immédiat, 360
Anneaux, 415
Antécédent,
    — et conséquent, voir conditionnel
    — et succédent, 156
Appartenance,
    relation d' —, 408
Application,
    — fonctionnelle, 297
Arbre,
    — binaire, 349
    — de preuve, 152
    — fini fermé, 245
    — fini ouvert, 245

— finitaire, 348
— indexé de formule, 243
— infini, 248
— infinitaire, 348
— sémantique, 347, 350
— sémantique complet, 351
— syntaxique, 107, 182
Arcs, 349
Arithmétique, 323
Assertion, 65
(formules) assertées, 65
asserter *versus* affirmer, 65
Assignation, voir interprétation
Atténuation, 151
Autonymie, 33
Axiomes, 108
— d'abstraction, 337
— du choix, 149, 426
— de séparation, 338, 406
— extralogiques, 309
indépendance des —, 119, 149
schémas d'—, 130
— *versus* hypothèses, 131
Axiomatisation,
axiomatisable, 86
axiomatisé, 86

Bivalence, 324, 328
principe de —, 32, 54
— *versus* logique floue, 98
— *versus* logique graduelle, 325
Borne,
— de complexité de l'instantiation et décidabilité, 252
Bouteille de Klein, 427

Calcul,
— qui se fait pas à pas, 82
*Calculus ratiocinator*, 40
Caractère, voir propriété, 173, 174
Carré logique,
— selon Euler, voir diagramme
— selon Venn, voir diagramme
Catégoricité, 311, 400
$\omega$- —, voir théorème de représentation, 401, 422
Catégorie,
définition de la —, 198
— grammaticale, 130
— syntaxique, 28, 106
Cercle vicieux, 29, 108
Champ, 415, 433
Chemin, 182, 264
Circonstanciels égocentriques, 64
Circuits électriques, 96–98, 428
Classe, 20
— complémentaire, 43, 59
— d'équivalence, 58
— élémentaire de modèles, 292
— universelle, 43
— vide, 43
— et univers, 48
Clause, 193
définition de —, 273
— fondamentale, 273, 347
Combinatoire, 95
Compacité,
— sémantique, 411
— syntaxique, 347
Complément, 46
Complétude,
— du calcul équationnel, 359
— "expérimentale", 76
— faible, 319
— forte, 319
— sémantique, 318
— syntaxique, 319
— syntaxique dans sa version sémantique, 320
— syntaxique dans sa version syntaxique, 320
— de la logique des propositions, 166
— de la logique des Stoïciens, 38
— et décidabilité, 89

— et énumérabilité, 88
— et récursivité, 86
Complexité,
— de la méthode des séquents, 179
— de la preuve par séquents, 180, 254
degré de — d'une formule, 241
Composition,
— de fonctions, 284
— de significations, 294
Concaténation, 297
Consension, 19
Conditionnel,
— diodoréen, 33
— philonien, 32, 55, 71, 78, 325
— vérifonctionnel, 37
introduction d'un —, 131
Conditions de vérité *versus* valeurs de vérité, 53
Congruence, 312
Connecteurs, 33, 168
— naturels *versus* formels, 95
pseudo- —, 327
Connexions, 181
Conséquence logique, 223, 314
— pour les formules fermées, 316
— pour les formules quelconques, 318
Consistant,
ensemble finiment —, 411
ensemble finiment — maximal, 411
schéma — (*satisfiable*), voir formule contingente, 99
Constante,
— -témoin, 413
définition inférentielle des — logiques, 168
Constituant, 22
définition de —, 106
— d'une formule, 214
— logique *versus* — syntaxique, 182, 238
Construction, 347
— grammaticale, 66
— d'un séquent, 163
la démonstration vue comme —, 32
Contenu,
— affirmé, 52
— d'information d'une théorie, 309
— *versus* forme, 28, 93, 331
Contexte,
— d'attitudes propositionnelles, 42
— de citation, 42
— extensionnel, voir principe d'extensionnalité
Contrainte,
— globale, 225, 232
— sur l'emploi de E.I. et G.U., 225
— sur l'unification, 282
Copule, 22
— interprétée en extension, 22, 23, 26
Corps, 415
Correction,
— du calcul des prédicats, 304
Correspondance bijective, 412
Couleurs,
théorème des quatre —, 426
Couple ordonné, 21
Coupure, 159, 165

Décision,
problème de la —, 81
Déclaratif,
— *versus* procédural, 152
Définition,
— explicite, 119
— implicite, 119

— inférentielle, 147
— non créative, 169, 170
— par récurrence, 416
— partielle, 333
Démonstration, 128
définition de la —, 128, 135
— la plus courte, 164, 167
— sans détour, 167
Démontrabilité,
— et vérité, 149, 336
Déploiement, 338
Dérivabilité, 326
Dérivation, 128
définition de la —, 143
— achevée, 227, 229
— globale, 143, 144
— locale, 143, 144
— normalisée, 146
Descendant, 360
Description définie, 68
Détachement, 115
Détour, 146, 167
Diagramme,
— d'Euler, 27
— de Venn, 48, 52
Direction argumentative, 63
Domaine,
— de référence, 74
— de Herbrand, 279, 346

Echange,
— dans le calcul des propositions, 120
— dans le calcul des prédicats, 209
Effectivité, 79
Enoncés,
— analytiques au sens de Frege, 404
— théoriques *versus* — d'observation, 76
Ensemble, 21
— finiment consistant, 411
— maximal, 411
— récursif, 416
— récursivement énumérable, 364,416
— vide, 21
famille d'—, 408
Enthymème, 61
Equations,
mise en — des syllogismes, 43
— *versus* identités, 205
Equipollents (axiomes —), 119
Equipotents,
modèles —, voir modèles
Equivalence, 123
relation d'—, 126
Equivoque,
faute de l'—, voir ambiguïté
Espace,
— de recherche, 91, 173, 195
— mémoire, 91
— topologique, 413
Espèce (au sens de Brouwer), 338
Etats de choses, 64
Evidence, 145
— et axiomes, 107
Evolution des espèces, 19
Expansion, 272, 420
Extension, 20
— conservatrice, 169, 327
— d'une théorie, 415
— élémentaire, 419
— *versus* intension, voir intension
Extensionnalité, voir principe d'—

Faits, voir états de choses
Fermeture transitive, 261
"*Fido*" - *Fido Theory*, 299
Finitaire *versus* infinitaire, voir arbres
Finitisme, 357
Foncteurs,
— de Skolem, 273

Fonction, 276
— caractéristique (logique), 84, 361
— de vérité, 32, 96
— récursive générale, 81
— récursive primitive, 83
composition de —, 83
limite d'une —, 74
Fonction propositionnelle, 35
— et proposition générale, 67
— et proposition singulière,66
Forme, 109
— clausale, 193
— fonctionnelle, 272
— grammaticale, 68
— implicative, 131, 139
— logique, 103, 104
— normale conjonctive (développée), 187
— normale disjonctive (développée), 187
— logique *versus* — contenu mathématique, 377
— propositionnelle valide, 101, 111
Formule, 79
— additionnelle (*versus* — principale), 158
— assertée, 136
— contingente, 385, 389
— moléculaire, 145
— principale, 158
— supposée, 136
Force logique, voir potentiel d'inférence
Fragment implicationnel, 116

Gardien,
— d'un programme, 367
Géométrie, 90, 323
Graphe,
— acyclique, 262, 265, 271
— cyclique, 272
Groupes, 89
— abéliens, 388
— abéliens sans torsion, 432
— involutifs, 388
— non abéliens, 383

Heuristique, 136
rôle — du dessin, 219
Homomorphisme, 297, 300
Hypothèse,
— auxiliaire, 131
— *versus* axiome, 129
preuve à partir d'—, 126

Identité,
— extensionnelle, 112
— intensionnelle, 112
— typographique, 112
— *versus* equation, 205
jugement d'—, 41
logique de l'—, 41
relation d'—, 41
Idéographie, 63, 64
Implication, 37
Impraticabilité, 91
Indépendance des axiomes, preuve d'—, 386
Indiscernabilité, 418
Induction,
— faible, 207
— forte, 207
— transfinie, 412
principe d'—, 80
Inférence immédiate, voir schéma d'—
Infini,
— dénombrable, 340
— indénombrable, 340
Instrumentalisme, 231
Intension, 20, 42, 96
— dans l'interprétation du calcul des propositions, 112
— des phrases, 53

— du prédicat "vrai", 307
— des syllogismes, 30
Interpolation,
théorème d'—, 428
Interprétation, 206, 224, 296
— de Herbrand, 343
— de Herbrand correspondante, 344
Intuition de l'espace, 219, 221
Intuitionniste, 338
Inversion,
principe d'—, 146
Isomorphisme, 300, 311, 400, 417
— *versus* équivalence, 418
Itération, 325, 368

Jugement, 19
— analytique *versus* synthétique, 404
— a priori *versus* a posteriori, 404
— d'existence, 59
— *versus* proposition, 65
— synthétique a priori, 404
classification des — en affirmatifs, négatifs, particuliers, universels, catégorique, hypothétiques, 65

Langue,
— artificielle, 62, 73
— formelle, 105
— naturelle, 62
— semi-formelle, 104
— universelle, voir idéographie, notation canonique, 75
Langage,
— *versus* théorie, 77, 293, 382
— -objet *versus* métalangage, 77, 78, 129, 142
Lemme,
— de König, 347
— du relèvement, 287
Lexique, 277
— autogénéré, 245
division du —, voir signature
Limite,
— d'une fonction, 74
*Lingua characteristica universalis*, 40
Littéraux, 187
Logique,
— des prédicats du premier ordre, 391
— du deuxième ordre, 272, 358
Lois,
— de Clavius, 362
— de dualité, 45
— d'exportation/importation, 123
— de substituabilité des identiques, 40
— de variation inverse de l'extension et de l'intension, 45

Machine,
— de Post, 365
Majorant, 409
Marque-place *versus* variables, 26, 34, 44
Mathématique *versus* métamathématique, 338
Matrice,
— et tableau matriciel, 96
— d'une formule en forme prénexe, 299
Maximal,
ensemble —, 411
Metalangage, 135
— sémantique, 78, 335
— syntaxique, 78, 335
variables du —, 131
Métathéorème,
— de la déduction en calcul des prédicats, 207

— de la déduction en calcul des propositions, 132, 153
— de l'échange en calcul des propositions, 120
— de l'échange en calcul des prédicats, 209
Métavariables, 131, 201
Méthodes,
— analytiques *versus* — synthétiques, 155
— axiomatiques *versus* — ensemblistes, 357
— axiomatiques pour fonder, 110
— pour décrire, 110
Minimalisation, 83
Minimalité, 406
Modalité, 331
Mode,
— synthétique, 165
— du raffinement, 165
Modèle,
définition de —, 301, 329
— de réalité, 381
— équipotent, 422
— fini *versus* — infini, 414
— non régulier *versus* structure, 384
sous- —, 419
Modularité, 225
*Modus ponens*, 34, 305
— et assertion, 65
— et détachement, 115
quatre formes du —, 127
Monotonie, 158, 252
Monoïde, 415
Moyen terme, 50
Multiplication étendue, 386

Négation, 121
Neutre, 46, 386
Noeud-échec,
— dans la méthode de Herbrand, 353
— dans la méthode des tableaux, 173
Noeud immmortel, 355
Nombre gödelien, 85, 87, 335
Normalisation,
— d'une preuve, 146, 168
— et induction mathématique, 237, 376
Notation canonique, voir langue semi-formelle, 63

Objet,
— intentionnel, 64
— physique, 75
— quelconque, 75
Occurrence, 183
— *versus* type, 40
Ontogenèse, 19
Opérateur,
— modal, 379
— *quatenus*, 66
Opération,
— *versus* relation, 56, 57
Opératoire, voir récurrence, 339
Ordre,
— partiel, 391
— total, 402, 409
"Ou",
— exclusif, 34, 37
— inclusif, 34, 37

Paradoxe,
— des ensembles, 337
— du menteur, 334
Paramètre, 211, 224 ,239
Parcours,
— en profondeur d'arbres, 182, 183
Paraphrase, 105
Partition, 384
Permutation, 254, 394

Pétition de principe, 29
Phrases,
— actives *versus* passives, 64
— intemporelles, 54
intension des —, 53
Plan de travail, 245, 247, 259
— de Kleene, 250
Pointeur, 261, 283
Polysyllogisme, 19, 124
Potentiel d'inférence, 64, 69
Prédication, 64, 198, 199
— double, 66
Présupposition, 47
— d'existence, 47, 71
Preuve, 86
— à partir d'hypothèses, 126, 127
— à partir d'hypothèses auxiliaires, 131
— de terminaison de programmes, 91
— formelle, 108, 109
— formelle et langue formelle, 109
— normale, 146
— par l'absurde, 362
Principe,
— d'abstraction, 337
— de la sous-formule, 166
— d'extensionnalité, 209
— *ex falso sequitur quodlibet*, 98
— d'héritage de propriété, 45
— d'induction, 80, 403, 405
— d'inversion, 146
Problème,
— du commencement, 18
— de correspondance, 368
dixième — de Hilbert, 431
instance d'un —, 365
réduction d'un — à un autre, 369
Productivité, 294
Programme, 17, 20
— d'une machine de Post, 366
extraction de —, 375, 376
Prolog, 273
Pronom,
— anaphorique, 74
— indéfini, 74
— personnel, 64
Proposition, 18, 22
— atomique *versus* moléculaire, 298
— attributive *versus* relationnelle, 67
— générale *versus* singulière, 67, 240
— hypothétique *versus* catégorique, 36,71
— primaire *versus* secondaire, 58, 59
— selon Platon, 22
— selon Aristote, 22
— selon Frege, 64
mise en équation des — fondamentales, 44
signification de — *versus* valeur de vérité de —, 208
Propriété, 21
— et intension, 21
Prototype, 45

Quantificateur, 69
— borné, 85
portée d'un —, 201
Quaternions, 434
Question expérimentale, 76

Raisonnement, 19, 25, 93, 104, 328
— en langue naturelle, 94
— informel, 94, 222
— guidé par l'intuition de l'espace, 219, 221
— mixte, 72
— par récurrence, 403

forme implicative d'un —, voir forme implicative
Rationalité, 232, 289
Récursivité, 86
Récurrence, 403
raisonnement par —, 339
Règles,
— de définition, 106
— de formation, 106, 109
— de transformation, 314
— d'inférence, 62
— facultatives, 138
— réversibles, 160
— structurales, 158
Régression à l'infini, 108, 406
Relation,
— d'appartenance, 408
— d'équivalence, 260, 312
— d'ordre, 260, 391, 408
— *versus* opération, 56, 57
Représentabilité, 415
Résolution, 193
Ruban de Möbius, 427

Satisfaction, 299
Saturation, 194, 280
Schéma,
— d'inférence, 33, 94, 142
Segments, 312
Semi-décidable, 373
Sens, 308
Séquents, 150, 255
Signature, 382
Signification, 90
Somme et produit arithmétiques, 44
Sophisme, 108
Sous-formule,
principe de la —, voir constituant, normalisation
Structure, 296, 384
— algébrique usuelle, 431
sous- —, 419
Substitution,
— en tant que règle d'inférence, 114
— exogène *versus* — endogène, 286
— structurale, 327
composition de —, voir unification
Substituabilité,
— des identiques, 40, 334
Supposition, 224
— *versus* thèse, 205
Surdétermination, 63
Syllogisme,
— catégorique, 25
— disjonctif, 34
— parfait, 39
Symétrique,
le —, 381, 386
Syntagme,
— nominal, 67
— verbal, 67

Tableaux,
— matriciels, 95
— sémantiques, 170
Tautologie, 57, 111, 210, 227, 385
Termes, 22, 56
— généraux *versus* singuliers, 20, 198
— généraux positifs *versus* généraux négatifs, 42
— indéterminés, 43
— libres pour une variable, 202
— logiques *versus* extralogiques, 330
Théorème,
— de Church, 375
— de compacité, 411
— de Craig, 423
— de Floyd, 373
— de Herbrand, 278, 352
— des quatre couleurs, 426
— de représentation, 313

— de Watanabe, 43
Théorie,
— axiomatisée *versus* axiomatisable, 86
— comme ensemble de thèses, 110
— des types, 358
— inductive, 421
— naïve des ensembles, 337
— d'ordre supérieur, 381
— physique, 76
extension d'une —, 415
inclusion d'une —, 415
Thèse,
— *versus* supposition, 205
Topic-comment, 64
Topologie, 423
Tore, 427
Type, 185
— *versus* occurence, 40
— primaire d'une formule *versus* — secondaire, 263

Unificateur principal, 289
Unification, 282
— et connexion, 259
— et méthode axiomatique, 288
— et résolution, 283
Univers du discours, 75

Vague, 98
Valeur de vérité,
— *versus* condition de vérité, 63, 110
Validité,
— *versus* vérité, 29
Variables, 100
— de classe, 55
— libres *versus* — liées, 201
— propositionnelles, 100, 103
— propositionnelles *versus* — individuelles, 333
métavariables (variables du métalangage), 120
interprétation conditionnelle des — *versus*, interprétation générique, 206, 224
Variante,
— d'une formule, 126
*x*- —, 300
Variation inverse,
loi de —, 45
Vérité, 18
— absolue *versus* — relative à un modèle, 77
— de correspondance, 306
— de logique, 37
— de mathématique, 77, 323
— d'Aristote, 305
— de Descartes, 306
— de Popper, 306
— de Putnam, 307
— de Tarski, 305
conception "sémantique" de la —, 307
définition récursive de la —, 300
degré de —, 325

CET OUVRAGE A ÉTÉ REPRODUIT
ET ACHEVÉ D'IMPRIMER PAR
L'IMPRIMERIE FLOCH À MAYENNE
EN NOVEMBRE 1997.

DÉPÔT LÉGAL : NOVEMBRE 1997.
Nº D'IMPRIMEUR : 42492.

*Imprimé en France*

## Langue – Raisonnement – Calcul

Présenté en deux volumes, *Logique* propose une mise au point complète sur l'utilisation de la logique, à partir des résultats et des méthodes d'exposition les plus récents. Le premier volume, *Méthodes pour l'informatique fondamentale*, contient un exposé des calculs des propositions et des prédicats, ainsi qu'une introduction aux théories du premier ordre. L'accent est mis sur l'aspect calculatoire, essentiel pour l'application en informatique et intelligence artificielle.

L'une des originalités du livre, premier manuel complet de logique en langue française, est de combiner le souci pratique avec une réelle rigueur conceptuelle. Cet ouvrage peut dès lors être un outil indispensable à tout approfondissement théorique.

Paul Gochet est professeur de logique et épistémologie à l'université de Liège. Il est membre de plusieurs sociétés savantes et a été titulaire de la chaire Francqui à l'université de Gand. On lui doit de nombreux travaux, notamment sur la philosophie de Quine.

Pascal Gribomont, docteur en sciences appliquées, est ingénieur au laboratoire de recherche Philips à Louvain-la-Neuve. Son travail concerne les applications de la logique à l'informatique, et plus spécifiquement à la programmation parallèle.

Editions HERMES
14, rue Lantiez
75017 Paris

9 782866 012496